

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Centro de Visualização e Optimização de Anúncios SAPO

José Leandro Amaro da Veiga

Mestrado Integrado em Engenharia Informática e Computação

Orientador: Luís Sarmento (FEUP)

11 de Julho de 2011

Centro de Visualização e Optimização de Anúncios SAPO

José Leandro Amaro da Veiga

Mestrado Integrado em Engenharia Informática e Computação

Aprovado em provas públicas pelo Júri:

Presidente: Jaime dos Santos Cardoso (Doutor)

Vogal Externo: Sérgio Guilherme Aleixo de Matos (Doutor)

Orientador: Luís António Diniz Fernandes de Moraes Sarmento (Doutor)

11 de Julho de 2011

Resumo

Numa Central de Anúncios em ambiente WEB onde a sua principal função está focada no ROI (Retorno do Investimento) dos seus anunciantes, é fundamental implementar soluções e ferramentas que possibilitem os seus utilizadores a bem gerir as suas campanhas. Esses utensílios deverão ser orientados aos objectivos das campanhas e destacam-se os relatórios completos de visualização gráfica e interactiva assim como ferramentas de optimização de anúncios/palavras-chave. Nessa perspectiva apresenta-se um conjunto de soluções e novos conceitos a aplicar no caso específico da Central de Anúncios do SAPO.

Os relatórios de *performance* das campanhas são imperativos em qualquer ferramenta orientada à WEB. Relatórios completos, com possibilidade de implementar condicionantes e restrições são implementados juntamente com gráficos visuais.

Assim, não só se complementa a informação estática apresentada anteriormente, como em certos casos se substitui por gráficos interactivos que permitam uma melhor percepção.

A optimização de campanhas corresponde a um conjunto de conceitos e ferramentas, orientadas na optimização individual de campanhas, que possibilitem ao anunciante melhor cumprir os seus objectivos aumentando o seu ROI.

Abstract

In a web environment, a Ads Center where its function is focused on the mainline Return of Investment of its advertisers, it is essential to implement solutions and tools that enable its users to manage their campaigns successfully.

These vessels should be towards the objectives of the campaigns and one should have full reports with complete graphics that could provide an interactive usage as well as optimization tools for both ads and keywords.

Agradecimentos

Agradeço ao ambicioso projecto SAPO LABs constituído pelo Laboratório da Universidade do Porto no departamento do DEEC.

Ao orientador Luís Sarmento, por ter aceitado orientar esta dissertação, assim como pela sua disponibilidade, profissionalismo e dedicação demonstrada. Muito aprendi com a sua frontalidade e competências.

Um agradecimento muito forte e especial para os meus pais e à minha namorada, que sempre prestaram o seu apoio em todos os meus desafios.

O Autor

Índice

1	Introdução.....	13
1.1	Contexto e Motivação.....	13
1.2	Definição do Problema.....	15
1.3	Restrições e Premissas.....	16
1.4	Objectivos.....	17
1.5	Estrutura da Dissertação.....	18
2	Gestão de Campanhas em Ambiente WEB.....	19
2.1.1	Introdução.....	19
2.2	Métricas de Análise de Negócio.....	19
2.2.1	Impression Share.....	19
2.2.2	Click Share.....	20
2.2.3	Conversion Share.....	22
2.2.4	Análise ao Anunciante AdvScore.....	22
2.2.5	Objectivos.....	24
2.3	Procedimento de Optimizações.....	25
2.3.1	Aumento do Número Tráfego.....	25
2.3.2	Aumento do CTR.....	25
2.3.3	Aumento do Retorno de Investimento (ROI).....	26
2.3.4	Optimização ao nível de Palavras-chave.....	26
2.3.5	Optimização ao nível de Anúncios.....	27
2.3.6	Optimização a Nível de Site.....	27
2.4	Descrição do Problema e Trabalhos Relacionados.....	28
2.4.1	Introdução.....	28
2.4.2	Descrição do Problema.....	29
2.4.3	Caracterização da Situação Actual do Anúncios SAPO.....	30
2.4.4	Conclusões.....	33
3	Trabalhos Relacionados.....	34
3.1	Modelos de Visualização.....	34
3.1.1	Polaris.....	34
3.1.2	Infovis.....	35
3.1.3	Protovis.....	35
3.1.4	Conclusão.....	36

3.2 Ferramentas de Auxílio de Negócio e Optimização.....	36
3.2.1 Sugestão de Palavras Chave.....	36
3.2.2 Optimização de Orçamento Publicitário (Budget).....	37
3.2.3 Estimativa de CTR com regras de decisão.....	38
3.2.4 Optimização a Nível do Website do anunciante.....	38
4 Desenvolvimento do Projecto.....	40
4.1 Estrutura do Documento.....	40
4.2 Pré-Processamento e Tratamento de Dados.....	40
4.2.1 Tipo de Dados.....	42
4.2.2 Critérios de Refinação.....	42
4.3 Relatórios Visuais.....	44
4.3.1 Informação Disponibilizada.....	44
4.3.2 Integração.....	45
4.3.3 Soluções Adoptadas.....	46
5 Implementação e Resultados Obtidos.....	47
5.1 Alterações às Base de Dados.....	47
5.2 Sistemas de pre-parsing.....	49
5.2.1 Visualização de Relatórios.....	49
5.2.2 Gráfico Impressões + Cliques + Conversões.....	49
5.2.3 Gráfico CTR com Conversion Rate e Impression Share.....	51
5.2.4 Gráfico Scatterplots Adg ID + Wrd ID.....	52
5.2.5 Optimizador.....	53
5.3 Análise de Resultados.....	54
5.4 Sugestões Automáticas de Optimização.....	55
5.5 Redefinição de Navegação.....	55
5.5.1 Campanhas.....	56
5.6 Necessidade de Processos de Cache.....	58
5.6.1 Introdução e Motivação.....	59
5.6.2 Resultados Obtidos.....	60
6 Conclusão e Trabalhos Futuros.....	62
6.1 Plano Técnico.....	62
6.2 Algoritmos/Contextualização.....	62
6.3 Conceitos e Negócio.....	63
6.4 Análise concorrencial relativa.....	64
6.5 Outros.....	65
6.6 Conclusões.....	65
7 Referências e Bibliografia.....	69
8 Anexos.....	71
8.1 Protovis: Exemplo de Aplicação.....	71
8.2 Frameworks/Ferramentas Instaladas.....	72
8.3 Tabelas de dados SQL.....	73
Índice Remissivo.....	79

Lista de Figuras

Figura 1: Resultados Patrocinados e Orgânicos.....	13
Figura 2: Interface desejada.....	18
Figura 3: Relação Volume com valor CPC.....	26
Figura 4: Níveis de optimização de contas.....	30
Figura 5: Relação Campanhas com Grupos de Anúncios com Anúncios.....	34
Figura 6: Relatórios do Anuncios SAPO.....	35
Figura 7: Estrutura e Relação Dados Brutos/Filtrados.....	43
Figura 8: Resumo de tabelas SQL.....	44
Figura 9: Diagrama de Venn para a Relevância.....	45
Figura 10: Interface Actual do Anúncios SAPO.....	47
Figura 11: Interface Desejada do Anúncios SAPO.....	48
Figura 12: Gráfico Linear Impressões, Cliques e Conversões (versão normal).....	51
Figura 13: Gráfico Barras de Impressões, Cliques e Conversões.....	52
Figura 14: Gráfico Linear com Condicionantes.....	52
Figura 15: Gráfico Linear com Condicionantes 02/11 a 16/11.....	52
Figura 16: Gráfico de Conversion Rate com CTR e Impression Share (versão normal).....	53
Figura 17: Gráfico de Conversion Rate com CTR e Impression Share (versão maior).....	54
Figura 18: Scatter plots com 10 keywords.....	54
Figura 19: Scatterplots + Piechart com 10 keywords.....	55
Figura 20: Esquema Actual de Navegação da secção Campanhas.....	58
Figura 21: Esquema Sugerido de Navegação da secção Campanhas.....	60
Figura 22: Sistema de Cache.....	62
Figura 23: Google Adwords - Análise concorrencial relativa.....	67
Figura 24: Exemplo de um gráfico/barras Protovis.....	75

Lista de Tabelas

Tabela 1: Análise global Impression-Share.....	21
Tabela 2: Análise global Clique-Share.....	22
Tabela 3: Opções de Matching da Indústria.....	28
Tabela 4: Passos propostos do algoritmo.....	40
Tabela 5: Estrutura Proposta Metas Temporais.....	65
Tabela 6: Tabela de Registo de Tempos de Execução cache/sem cache no Gráfico #1.....	69
Tabela 7: Tabela de Registo de Tempos de Execução cache/sem cache no Gráfico #2.....	69
Tabela 8: Tabela de Registo de Tempos de Execução cache/sem cache no Gráfico #3.....	70
Tabela 9: Tabela SQL adgroup_word.....	82
Tabela 10: Tabela SQL adg_adgroups.....	82
Tabela 11: Tabela SQL Adsense_AffiliateSite.....	83
Tabela 12: Tabela SQL Adsense_PositionInSite.....	83
Tabela 13: Tabela SQL ads_ads.....	84
Tabela 14: Tabela SQL bid_impressions.....	84
Tabela 15: Tabela SQL cam_campaigns.....	85
Tabela 16: Tabela SQL clicks.....	85
Tabela 17: Tabela SQL conversions.....	85
Tabela 18: Tabela SQL site_impressions.....	86
Tabela 19: Tabela SQL usr_impressions_report.....	86
Tabela 20: Tabela SQL wrd_adg_performance_report.....	87
Tabela 21: Tabela SQL wrd_performance_report.....	87

Abreviaturas e Símbolos

WWW	<i>World Wide Web</i>
Afiliado	Qualquer pessoa física ou jurídica poderá ser um Afiliado. Um Afiliado é uma pessoa/empresa que difunde campanhas de publicidade no seu website em troca de comissões.
Anunciante	Um anunciante é uma empresa (ou individual) que paga para anunciar os seus produtos, serviços ou website na plataforma de afiliados. O objectivo de um anunciante é garantir audiência e potenciais clientes para os seus produtos/serviços. Cabe aos afiliados a responsabilidade de difundir as campanhas, ganhando diferentes comissões de acordo com os objectivos que atingem.
Banner	Um banner é a forma publicitária mais comum na internet. É um sistema muito usado para a divulgação de websites na internet, divulgar produtos/serviços, entre outros. Os banners são geralmente imagens desenvolvidas em formato .jpg ou .gif, mas também podem ser animações em Java, Shockwave ou Flash. Um banner poderá ter diferentes dimensões, no entanto, existem algumas dimensões padrão. Geralmente, as imagens possuem uma forma alongada, vertical ou horizontalmente.
Campanha	Uma campanha online consiste na difusão de um ou mais anúncios de texto/banners que são exibidos em diversos websites. Geralmente, as campanhas online são criadas para atingir um objectivo de marketing específico (por exemplo, uma campanha de promoção de verão, lançamento de um produto, divulgar um serviço, entre outros). Normalmente, uma campanha possui sempre um período e um orçamento associado, que poderá ser renovado sempre que o anunciante o pretender. Os Afiliados por sua vez, vão divulgar as campanhas online nos seus websites, ganhando comissões sobre os objectivos que atingem.
Clique	Cada clique efectuado num link de texto ou banner poderá ter um valor associado. O clique possui assim, um valor de comissão por cada clique válido realizado.
Cookie	Um cookie é um pequeno texto que os sites podem enviar aos navegadores, anexado a qualquer conexão. Nas visitas posteriores o navegador reenvia os dados para o servidor dono do cookie.
CPL	O CPL (Cost Per Lead) é um termo usado para definir o custo por Lead/Formulário? . É o valor que um anunciante paga cada vez que um

CPM	<p>utilizador realiza uma acção ou preenche um formulário.</p> <p>O CPM (Cost Per thousand impressions) é um termo usado para definir o custo por cada Mil impressões. É o valor que um anunciante paga a um website, pela exibição de um banner publicitário por cada mil impressões realizadas.</p>
CTR	<p>O CTR (Click Through Rate) é uma taxa utilizada para calcular a percentagem de cliques em um anúncio, a dividir pelo número de impressões, multiplicando o valor por 100. Por exemplo, um CTR de 2%, indica que a cada 100 impressões do seu anúncio, somente 2 anúncios foram clicados.</p>
Duplo-Clique	<p>O Duplo-Clique é contabilizado quando um utilizador clica num determinado anúncio e posteriormente realiza um novo clique no site do anunciante, contabilizando 2 cliques. É o valor que um anunciante paga a um afiliado quando um utilizador realiza os 2 cliques, um no anúncio e outro no website do anunciante.</p>
CPC	<p>Custo efectivo por clique. O custo de uma campanha, dividido pelo número total de cliques emitido por essa campanha. Calcula-se da seguinte forma:</p>
FAQ	<p>Total Valor Pago / Total de Cliques Recebidos.</p> <p>FAQ é um acrónimo da expressão inglesa Frequently Asked Questions, que significa Perguntas Frequentes. Um FAQ, quando usado num contexto pluralista, significa uma compilação de perguntas frequentes acerca de determinado tema. Quando usado num contexto singular, um FAQ será uma dessas perguntas frequentes.</p>
GIF (Graphic Interchange Format)	<p>Um GIF (Graphics Interchange Format) é um formato de imagem de mapa de bits muito usado na world wide web, quer para imagens fixas, como para animações.</p>
HTML	<p>HTML (acrónimo para a expressão inglesa HyperText Markup Language, que significa Linguagem de Marcação de Hipertexto) é uma linguagem de marcação utilizada para produzir páginas na Web. Os documentos HTML são interpretados pelos navegadores/browsers.</p>
Hiperligação	<p>Uma Hiperligação (Link) é uma referência num website a outras páginas do mesmo site, ou a outro website externo. Pode ser acedido através de uma palavra, frase ou imagem num website.</p>
Impressão	<p>A visualização de uma página ou anúncio pelo utilizador. Cada vez que um website é carregado, origina uma impressão.</p>
Lead	<p>Uma Lead é uma informação de contacto submetida, uma acção específica como subscrever uma newsletter, ou equivalente por um visitante de um website. As empresas que não vendem online, podem utilizar o seu website para gerar leads.</p>
Newsletter	<p>Uma Newsletter é um tipo de publicação de distribuição regular a assinantes e que aborda geralmente um determinado assunto. Geralmente, um utilizador recebe uma newsletter sempre que submete o seu email num website a pedir o envio da mesma, ou após efectuar um registo num determinado website.</p>
PPC	<p>PPC são as iniciais de Pay Per Click, em português, paga por clique. É a forma mais rápida de gerar visitantes qualificados para um determinado</p>

	website, através da colocação de anúncios textuais ou gráficos nos motores de busca e websites.
Programa de Afiliados	Um Programa de Afiliados é um tipo de marketing feito por websites/empresas que contratam pessoas à comissão para que estes possam divulgar o seu website, produtos/serviços. Os Webmasters ao efectuarem o registo num Programa de Afiliados, poderão divulgar links, banners e outro tipo de propaganda publicitária, recebendo uma comissão pelos resultados. É um tipo de marketing de resultados, onde o Webmaster, apenas recebe comissões por cliques, impressões, leads/formulários preenchidos, vendas.
Taxa Conversão	A taxa de conversão é a relação dos visitantes que convertem visualizações em acções com base na solicitação de um anunciante. Calcula-se da seguinte forma: $\text{Total Conversões} / \text{Total Visualizações} \times 100$
Testes A/B	Fazer bastantes testes nas campanhas que promove é uma excelente forma de aumentar os seus rendimentos. Através dos testes A/B, o afiliado pode comparar duas campanhas em simultâneo, mostrando as mesmas de forma alternada, de forma a poder avaliar melhor os resultados.
URL	Um URL (Uniform Resource Locator), em português, Localizador-Padrão de Recursos, é o endereço de um recurso (um arquivo, uma impressora, etc), disponível em uma rede; seja a internet, uma rede corporativa, uma intranet. Exemplo: http://www.website.com/seccao/afiliados.html

1. Introdução

1.1. Contexto e Motivação

Com o crescimento da Internet e com a sua consolidação como um novo meio de negócios, existiu a necessidade de implementar ferramentas e utensílios que permitam às empresas obter rendimentos dos seus sítios (*sites*) da internet. Em concreto, deverão ser implementadas soluções que permitam aos sítios populares ou fornecedores de tráfego, tirar partido financeiro da sua audiência, rentabilizando-a com anúncios de terceiros.

The image shows a screenshot of a search engine results page (SERP) for the query "carros usados" on the sape.pt website. The page is divided into two main sections, each highlighted with an orange box and an arrow pointing to a label on the right.

Top Section (Sponsored Results): This section is labeled "Resultados Patrocinados" (Sponsored Results). It contains three search results, each with a small icon, a title, and a brief description. The first result is "Classificados Auto Sape" from auto.sape.pt, with the text "Venda já o seu carro! O anúncio é Grátis!". The second result is "Nº1 em Carros Usados" from www.standvirtual.com/carros-usados, with the text "O StandVirtual® é o Nº1 em Portugal Comprar/vender Só no StandVirtual®!". The third result is "Carros Usados" from www.custogusto.pt/anuncios-gratis, with the text "Venda aqui o seu carro. A Custo Zero!".

Bottom Section (Organic Results): This section is labeled "Resultados Orgânicos" (Organic Results). It contains several search results, each with a small icon, a title, and a brief description. The first result is "Carros Usados Portugal" with the text "Anúncios de carros usados em Portugal. Pode anunciar gratuitamente o seu automóvel usado nos nossos classificados. Se quer comprar ou vender automóveis então visite-nos e ...". The second result is "Carros Usados, Stand Virtual, Venda Automóveis Usados" with the text "Carros usados, Stand virtual, Venda de Automóveis usados. Carros usados baratos só no nosso Stand Virtual, carros semi-novos, novos, todos os preços.". The third result is "Carros Usados, Automóveis Usados, Venda de Carros - Novo Dono" with the text "Anúncios de carros usados e auto usados, vendas de carros e classificados de carros. Motos Usadas e Barcos Usados.". The fourth result is "Carros Usados, Automóveis Usados, Classificados Auto - Autoportal" with the text "Carros usados, Automóveis Usados e Carros Novos em Portugal. A inserção de Carros Usados está disponível para Profissionais e Particulares.". The fifth result is "Carros Usados de Profissionais - Automóveis-Online" with the text "Carros usados, novos e semi-novos. Compre carros usados com segurança só a profissionais.". The sixth result is "Carros Usados, Automóveis, Motos e Barcos - AutoOcasão" with the text "Carros Usados e Automóveis à venda. Comprar carros usados, motos usadas, peças, autocaravana, caravana. Anúncios de carros usados Portugal aqui!".

Figura 1: Resultados Patrocinados e Orgânicos

O portal SAPO é um conjunto de sítios e serviços pertencentes à empresa PT Multimedia.com, Serviços de Acesso à Internet SGPS, (também conhecido como PTM.COM) que tem centenas de milhar de visitantes por dia. O Anúncios SAPO é a ferramenta que permite à PTM.COM rentabilizar essa audiência, criando um vínculo entre anunciantes terceiros e os seus vários sítios. Um anunciante que seja utilizador do Anúncios SAPO, cria uma conta, cujos anúncios tanto podem ser em formato texto ou formato de imagem, e indicam um conjunto de palavras-chave ou URL's onde desejam que os mesmos sejam apresentados.

Assim, a cada pesquisa efectuada no Portal SAPO, é devolvido ao utilizador os chamados resultados orgânicos, constituídos pelos sítios propriamente ditos que se aproximam do termo de pesquisa, e os resultados patrocinados, compostos por anúncios de anunciantes.

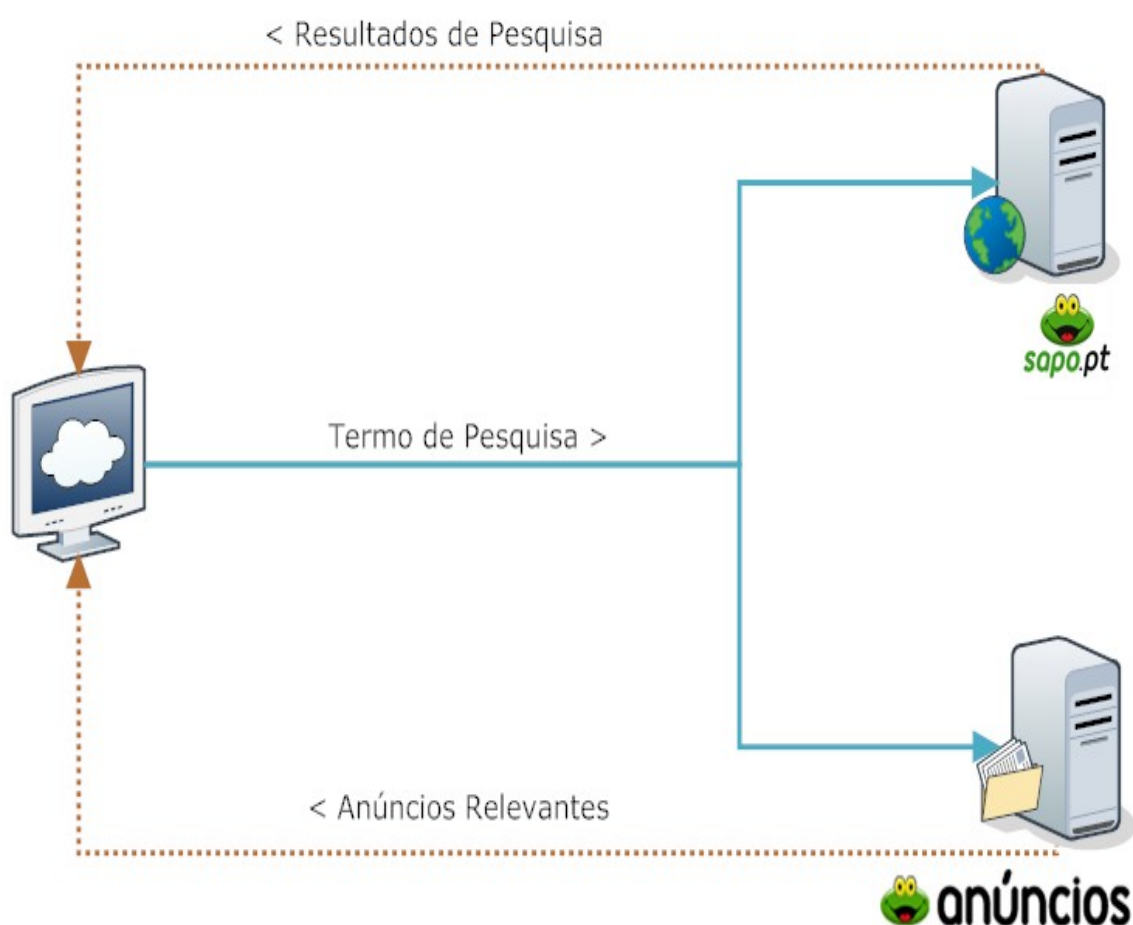


Ilustração 1.1: Relação anúncios patrocinados e orgânicos

1.2. Definição do Problema

Toda a gestão de anúncios, como criação, edição, remoção e gestão de palavras-chave, é efectuada numa zona privada, onde só os anunciantes têm acesso (protegido por um nome de utilizador e password) que tem como nome de “Central de Anúncios SAPO”¹. É nessa central que o anunciante gere os seus anúncios, obtém relatórios e efectua carregamentos monetários de conta. A versão actual da “Central de Anúncios SAPO” é estável e funcional, mas contém algumas limitações:

- Com a informação das campanhas são apresentados alguns dados de “*performance*” das mesmas, mas não existe flexibilidade de as definir. O utilizador não está possibilitado de definir restrições temporais nem que dados pode consultar.
- A secção de relatórios é restrita a impressões ou “pageviews” (número de vistas de um anúncio), “cliques” (número de cliques de um anúncio), “CTR” (relação vistas de um anúncio com cliques) e “conversões” (objectivos obtidos: venda, preenchimento de formulário, etc).
- Não há possibilidade de consultar dados geográficos (por região/país).
- A sobreposição de dados estatísticos é inexistente. Por exemplo, não é possível relacionar o “CTR” com a taxa de conversões.
- Completa inexistência de relatórios visuais a acompanhar a informação estática.

Os anunciantes, baseados em alguns inquéritos efectuados, consideram a actual plataforma insuficiente para o aumento do ROI (Return Of Investment)² das suas

¹ A Central de Anuncios SAPO pode ser consultada pela internet em <http://anuncios.sapo.pt>

² Return Of Investment, ou Retorno de Investimento corresponde ao valor efectivo de retorno de um investimento publicitário, geralmente expresso em percentagem. Por exemplo, um ROI de 120% significa que numa campanha de marketing, o anunciante recuperou 120% do seu dinheiro investido.

campanhas, tendo que depender de ferramentas externas para a compilação de relatórios como, por exemplo, o Excel. Uma das funcionalidades mais requeridas pelos anunciantes é a elaboração de gráficos intuitivos que lhe permitam de forma visual encontrar as mais, e as menos valias das suas campanhas. Os relatórios visuais deverão ser o mais intuitivo possível e os seus utilizadores deverão ter acesso à informação de forma instantânea e directa.

O desenvolvimento da interface utilizador, será efectuado com o auxílio de APIs JavaScript como jQUERY que em conjunto, contribuirá para uma interactividade total. Face à natureza de dados em formato “bruto”, será necessária a filtragem e pré-processamento de dados. Para esse objectivo será necessário recorrer a processamento da linguagem natural com expressões regulares.

De salientar que não haverá qualquer alteração aos dados, isto é, o acesso aos dados será sempre em modo de consulta mantendo-se os mesmos sempre inalterados³.

Mesmo o próprio pré-processamento de informação não implicará qualquer alteração aos dados originais.

1.3. Restrições e Premissas

A nova versão dos relatórios e zona de optimização do “Anúncios SAPO”, é desenvolvida em ambiente Linux, utilizado PHP na versão 5.3.

Actualmente os dados são armazenados em “MySQL” no motor “InnoDB” e fazem uso de diversas linguagens de programação como Python, PHP, Perl entre outras.

A nova versão tira partido das condições actuais, pois entendem-se como adequadas para o desenvolvimento do projecto. O PHP é uma linguagem bastante actual e suficiente, e o motor InnoDB é adequado pois garante integridade dos dados e apresenta uma robustez suficiente que pode a qualquer momento ser escalada com as tecnologias de computação distribuída já existente no SAPO.

³ No presente projecto, alterou-se de facto a estrutura de tabelas de SQL e novos campos, mas mesmo assim, os dados foram preservados.

1.4. Objectivos

Os relatórios visuais deverão ser o mais intuitivos possível com o máximo de interactividade. Não só se pretende disponibilizar os dados de forma gráfica e fácil de interpretar, como também se pretende dar os dados certos ao anunciante de forma a obter o máximo de ROI. De notar que a oferta de dados deverá ser futuramente configurável, pois nesta área rapidamente se encontram novas fórmulas, novos conceitos para além dos apresentados, e claro, novas ideias.

Assim, o serviço a disponibilizar deverá ter como prioridade a facilidade e versatilidade. Com isto em mente e tendo em consideração o feedback dos anunciantes, o projecto terá como prioridades:

- Pré-processamento de dados brutos com a eliminação dos dados redundantes ou irrelevantes para a análise/gestão;
- Interpretação e escolha dos dados mais adequados com vista a um aumento do ROI para disponibilizar aos anunciantes;
- O actual sistema faz uma análise univariada, pelo que será prudente estabelecer análises multivariadas que permita a análise simultânea de múltiplos dados;
- O sistema deverá ser o mais interactivo possível.

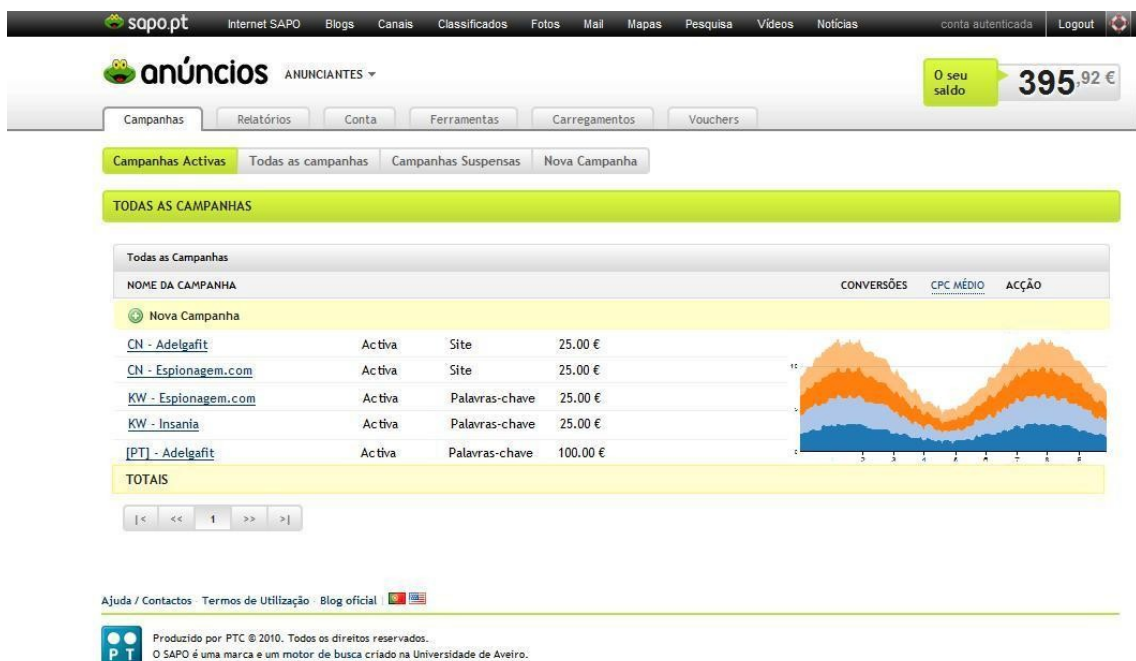


Figura 2: Interface desejada

O protótipo da interface será o semelhante ao da Figura 2 (17), onde há uma sintonia entre a informação textual e gráfica. O objectivo será manter uma interface familiar ao anunciante, completando em certos casos com gráficos, e ao mesmo tempo desenvolver uma nova área de relatórios/acompanhamento que permitam uma visão mais global.

Neste tipo de campanhas é essencial ao anunciante poder analisar com exactidão o impacto das alterações às suas campanhas, pelo que a interface deverá ser rica em ferramentas comparativas. Deverão ser implementadas comparações temporais, antes/depois/durante as alterações, mais conhecidos como dados de testes A/B⁴.

1.5. Estrutura da Dissertação

O primeiro capítulo contém uma introdução ao trabalho com informação relevante em relação ao contexto e objectivos, assim como conceitos necessários à compreensão do documento e projecto.

O segundo capítulo contém os Trabalhos Relacionados, que identificam o trabalho já desenvolvido na área, quer a nível empresarial, quer a nível académico. São também apresentadas possíveis soluções aos problemas apresentados no capítulo anterior.

O terceiro capítulo contém o desenvolvimento do projecto e a maneira desejada de implementar as soluções apontadas.

O quarto capítulo apresenta a implementação propriamente dita, nomeadamente as alterações efectuadas, novidades implementadas e conclui-se com uma análise crítica aos resultados obtidos.

Por fim, no quinto e último capítulo é feita uma conclusão de toda a dissertação, uma análise a todo o funcionamento da Central de Anúncios, como um todo, e são também definidos os trabalhos futuros.

⁴ O Google apresenta uma definição de testes A/B em [<http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=71976>] (16/06/2011)

2. Gestão de Campanhas em Ambiente WEB

2.1.1 Introdução

A publicidade online na Internet é hoje em dia considerada de importância absoluta para qualquer empresa com uma presença online. Seja uma empresa de pequena ou grande dimensão ou uma empresa cujo *core-business* seja inteiramente dedicada à Internet, todas elas, deverão poder utilizar os mesmos procedimentos para criar e gerir as suas campanhas. Uma plataforma de anúncios deverá, assim, permitir a qualquer tipo de entidade, independentemente do seu *know-how*, oferecer todos os mecanismos necessários.[JGPPC]

2.2. Métricas de Análise de Negócio

2.2.1 Impression Share

O Impression Share é uma métrica que representa a percentagem de visualizações de um anúncio que uma determinada palavra-chave obteve, relacionando todas as visualizações de anúncios de concorrentes para essa mesma palavra-chave. Assim, por exemplo, uma campanha que tenha uma impression-share de 25%, significa que só teve os seus anúncios apresentados $\frac{1}{4}$ das vezes.

O impression-share tem como principal objectivo, ao fornecer o desempenho relativo na rede, apresentar a fatia inexplorada da rede. Pode assim fazer acertos aos seus valores de CPC e optimização a nível de conta, de forma a garantir o máximo proveito de toda a rede.[ISLOST]

$$f_{ik} = \frac{n_{ik}}{\sum_{ik} n} \quad (2.1)$$

$$f_i = \frac{n_i}{\sum_i n} \quad (2.2)$$

Para calcular esta métrica, para todo o i faz-se corresponder uma palavra-chave n e k uma campanha.. Podemos assim calcular para toda a palavra-chave o impression-share. (2.1).

De igual modo, podemos calcular o impression-share para toda a rede. (2.2)

Pelos dados fornecidos, no período de trabalho, podemos aferir que estiveram em rotação cerca de 112.231 palavras-chave, e cerca de 272.015 palavras a competir a essas mesmas campanhas, dando assim uma conclusão de:

N.º Palavas-chave distintas	112231
Campanhas concorrentes	272015
Média palavras-chave por campanha	2,43 / campanha
Média de Impression-share	37.8%

Tabela 1: Análise global Impression-Share

Uma média de impression-share alta significa que existe pouca concorrência. Neste caso, podemos deduzir que existe uma média de 37.8%, ou seja, pelo impression-share podemos deduzir que existe em média 3 concorrentes para uma mesma palavra-chave.

Como este valor se aproxima do número médio de palavras-chave por campanha, podemos concluir que a rede apresenta um valor saudável de concorrência, estando próximo da sua rentabilização máxima.

2.2.2 Click Share

O Click Share, semelhante ao impression-share mencionado acima, é uma métrica que representa a percentagem de cliques de um anúncio que uma determinada palavra-chave obteve, de todas os cliques de anúncios de concorrentes para essa mesma palavra-chave. Assim, atendendo a uma campanha que tem um click-share de 50%, significa isto que essa mesma campanha só teve os seus anúncios clicados metade das vezes.

O click-share tem como principal objectivo, ao fornecer o desempenho relativo na rede, apresentar a fatia pouco explorada da rede. O click-share ajuda a identificar com eficácia anunciantes que ocupam um grande volume de visualizações, mas que no seu final possuem fraco CTR, resultando em pouca margem comercial.

$$f_{ik} = \frac{n_{ik}}{\sum_{ik} n} \quad (2.3)$$

$$f_i = \frac{n_i}{\sum_i n} \quad (2.4)$$

Para ser calculado o click-share, onde i corresponde a uma palavra-chave n e k corresponde a uma campanha, podemos calcular para toda a palavra-chave o click-share. (2.3).

De igual modo, podemos calcular o click-share para toda a rede. (2.4)

Pelos dados fornecidos, no período de trabalho, podemos aferir que se obtiveram cliques em cerca de 79,109 palavras-chave, e cerca de 8,172 campanhas a competir a essas mesmas palavras-chave, dando assim uma conclusão de:

N.º Palavras-chave distintas clicadas	79109
Campanhas concorrentes	8172
Média palavras-chave por campanha	9,68 / campanha
Média de Click-share	84,29%
Média de CTR	0,01%
Média de CPC	0,11 €

Tabela 2: Análise global Click-Share

Com os dados da rede, podemos verificar se um determinado anunciante está com a sua campanha optimizada. Neste caso, uma campanha pouco optimizada significa que os seus anuncios são pouco relevantes. Isto porque as palavras-chave têm licitações significativas para ser apresentadas na rede, mas não estão a apresentar cliques. Como os cliques são o último objectivo, deverá ser tido em conta tais dados, tendo em atenção a sua rentabilização.[PRFING]

Assim, por estes dados podemos concluir que o click-share, em média, é alto, o que significa que o volume de tráfego está concentrado em poucos anunciantes. Assim, até aqui, pela análise das métricas de impression-share, conclui-se que a rede está devidamente explorada mas pouco optimizada para o rendimento monetário ao criar pouco volume de clicks. Temos em cada 10.000 visualizações de um anúncio, uma média de um clique. Traduz-se, em média, no pagamento de € 0,11 por parte dos anunciantes por cada mil visualizações. É um valor muito abaixo da média, pois temos como referência o valor de 2% anunciados pelo Google Adwords ⁵[GACR2]. Podemos deduzir que o SAPO, mediante os dados em estudo, apresenta um valor duzentas vezes inferior ao do seu maior concorrente, o Google Adwords.

2.2.3 Conversion Share

O Conversion Share, semelhante ao click-share mencionado acima, é uma métrica que representa a percentagem de conversões de um anúncio que uma determinada palavra-chave obteve, de todas as conversões geradas dos concorrentes para essa mesma palavra-chave. Assim, atendendo a uma campanha que tem um conversion-share de 75%, significa isto que essa mesma campanha obteve $\frac{3}{4}$ de todas as conversões.

O conversion-share ajuda a identificar, com eficácia, os anunciantes que obtenham mais resultados, e mediante estes dados, tendo embora em consideração que nem todos os anunciantes usam actualmente o *tracking* de conversões, podemos migrar o sistema para um modelo de pagamento à conversão, em vez de pagamento ao clique.

$$f_{ik} = \frac{n_{ik}}{\sum_{ik} n} \quad (2.5)$$

$$f_i = \frac{n_i}{\sum_i n} \quad (2.6)$$

⁵ A maior plataforma de anúncios é actualmente o Google Adwords, mas nunca relevou publicamente os dados para Portugal. Sendo assim, considerando o SAPO como o número dois, deveremos hipoteticamente considerá-los como os valores médios da indústria, baseando esta comparação na rede global a ter em conta.

Para ser calculado o conversion-share, onde i corresponde uma palavra-chave n e k corresponde a uma campanha, podemos calcular para toda a palavra-chave o conversion-share. (2.5).

De igual modo, podemos calcular o conversion-share para toda a rede. (2.6) Esta métrica será de fundamental análise quando houver, e se houver, uma migração para um sistema de PPA – Pay Per Action.

2.2.4 Análise ao Anunciante AdvScore

Tendo em atenção todas as métricas, podemos fornecer um “peso” a cada uma delas, e mediante o historial de um anunciante podemos aferir uma pontuação. No caso, cada anunciante possui um AdvScore que consiste numa pontuação de 0 a 100, em que 100 será a nota do melhor anunciante da rede. Isto é, existe obrigatoriamente pelo menos um anunciante com nota 100, e obrigatoriamente um anunciante com nota 0 correspondendo ao melhor e pior, respectivamente.

As métricas aqui apresentadas serão insuficientes para chegar a uma pontuação, por isso deveremos considerar certos parâmetros que convém aplicar, pelo que se descrevem:

- Historial de Pagamentos [FIDELIDADE=pag]– Partindo do lato princípio que um anunciante fiel utiliza a plataforma continuamente porque encontrou um equilíbrio entre montante gasto e montante de retorno, vamos “premiar” aquele anunciante que efectua pagamentos recorrentes. Assim, faz-se corresponder este factor como **20%** da nossa escala. O algoritmo deverá fornecer 20 pontos ao anunciante que mais pagou, e 0 pontos ao que menos pagou.
- Ética/Políticas Empresariais [FIDELIDADE=pol] – Sempre que um anunciante cumpra as regras e utilize sites fidedignos, vamos usar um peso de **10%** na nossa escala. Assim, qualquer anunciante tem por defeito 10 pontos, que poderão ser reduzidos a 5 se forem sites para adultos, por exemplo, 2 se forem jogos online ou outros conteúdos, etc. Esta escala deverá ser subjectiva e ter em conta os objectivos da rede. Parte-se do princípio que o tipo de anunciantes e quem eles são, é importante, pelo que se considera 10% um peso suficiente.
- Relatividade [PERFORMANCE=rel] – Um bom anunciante é aquele que no seu conjunto, tenha *performance* igual ou superior à rede, o que se traduz em mais rendimentos para o próprio SAPO. Assim, cada anunciante que cumprindo as políticas próprias e internas do SAPO e tenha um CTR mais alto que a média, deverá ser considerado um bom anunciante. Assim, oferecemos um peso de **35%** para este valor. Se um afiliado conseguir ter o melhor CTR da rede deverá ter 35 pontos pois tem um comportamento, que a longo-prazo, se traduz em

maior rendimento para o SAPO. Este valor só se aplica, se evidentemente as campanhas estiverem em CPC.

- Métricas “share” [PERFORMANCE=share] – As métricas de Share deverão ser consideradas com um valor de **25%**. Entende-se que são importantes, mas não tão importantes como a Relatividade. Quem tem mais share deverá assim continuar pois demonstra ser um bom anunciante.

Pelo exposto, tendo em consideração os valores aqui definidos, podemos chegar a uma equação que, para cada anunciante, é assim calculada:

$$AdvScore_a = (pag_a * 0,2) + (po_a * 0,10) + (rel_a * 0,35) + (share_a * 0,25)$$

equation 2.1: Cálculo de AdvScore

Para os novos anunciantes deverá ser considerado uma pontuação inicial neutra, ou seja de 50%. Só assim se garante igualdades aos novos anunciantes, dando oportunidade, dentro dos limites, de competir com anunciantes mais valiosos ou que estejam a anunciar há mais tempo.

Estas métricas e algoritmos deverão estar em constante refinamento e só se deverá considerar uma fórmula funcional (e isto sempre temporária), após alguns dias de uso.

2.2.5 Objectivos

Cabe ao anunciante estabelecer os seus objectivos quando gere uma campanha online. O pré-planeamento é essencial e a estratégia deverá estar previamente definida antes de criar uma campanha. Assim, o próprio anunciante deverá ter em conta as suas margens, definir um orçamento limite, e essencialmente, qual a percentagem da sua margem estaria disponível a partilhar com a plataforma de anúncios. Em regra, tendo em conta o sistema de licitações em vigor, podemos concluir que quanto maior for o montante que estiver disponível a pagar, maior será o volume gerado, mas menor será a sua margem de lucro.[PUBL4]

O anunciante deverá ter que criar a campanha com esses factores em mente, tomando desde já a sua primeira decisão de negócio.

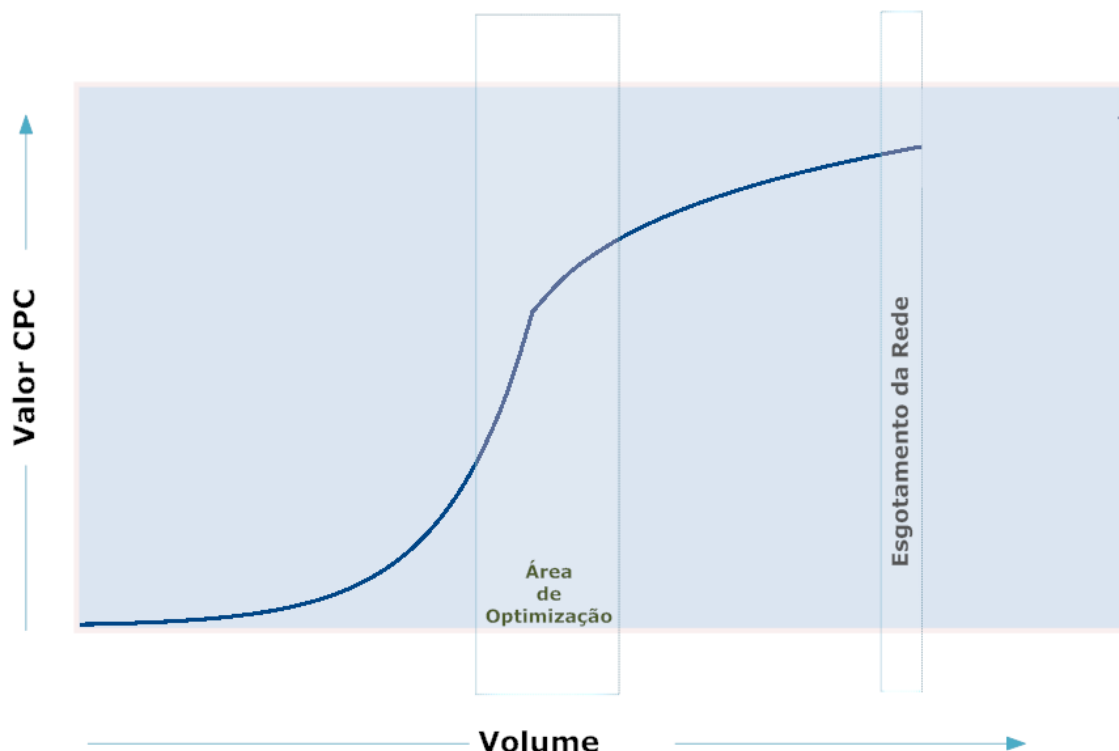


Figura 3: Relação Volume com valor CPC

É possível assim concluir facilmente que em negócios de um determinado tema, existe a chamada área de otimização, que corresponde, em regra, à média de valores CPC que os anunciantes desse determinado nicho estão disposto a pagar (chegando assim ao valor relativo de uma determinada CPC e verdadeiro valor de um nicho). Podemos verificar que a Área de Otimização de um nicho rentável e competitivo, como o do ramo automóvel, estará situada em valores médios de CPC bastante superiores que outros nichos menos competitivos, como por exemplo o nicho da educação.⁶

A tendencia do valor de CPC é então crescente até chegar à área de otimização, pois existe bastante concorrência, e depois estagna nessa mesma área até ser atingido o esgotamento de tráfego. Quer isto dizer que o tráfego é todo adjudicado por esses valores de CPC.

2.3. Procedimento de Otimizações

Tecnicamente deverá ser elaborado um conjunto de acções para cada estratégia de optimização. Essas acções serão enumeradas no próximo capítulo.[OPTKenv]

Importa mencionar os estudos internos do próprio Google em que apresentam casos reais com evoluções significativas dos resultados dos anunciantes em diversos temas como no retalho, [CASEST] ou viagens[CASETR].

Uma análise meticulosa de todos os procedimentos de optimização disponíveis, são a chave para uma campanha de sucesso, duradoura, e rentável tanto para o anunciante como para a própria plataforma.[GGICASE]

⁶ O estudo de nichos e mercados, por si só, seriam motivo suficiente de criação de uma dissertação própria.

2.3.1 Aumento do Número Tráfego

Neste caso, a otimização corresponde ao número de palavras-chave a correr na rede. Deverão ser sugeridas, por exemplo, novas palavras-chave baseadas nas já existentes, que poderão consistir em sinónimos, singular/plural, termos equivalentes e outros. A otimização deverá focar-se claramente no aumento de keywords.

2.3.2 Aumento do CTR

A otimização de um CTR corresponde a um aumento de cliques com as mesmas visualizações. Assim, deverão ser otimizados os anuncios na sua qualidade melhorando a relevância em relação às pesquisas efectuadas. A otimização deverá ser focada em relacionar o mais possível entre as palavras-chave e os anúncios. Para além disso, as que são consideradas irrelevantes ou que desvirtuem o resultado, deverão ser consideradas como “Negativas”.

$$CTR = (N.^{\circ}Clicks / N.^{\circ}Pageviews) * 100$$

equation 2.2: Cálculo de CTR

2.3.3 Aumento do Retorno de Investimento (ROI)

A otimização é equivalente à otimização do Aumento do CTR com o aumento da relevância das palavras-chave com os anuncios, mas também com o próprio site do anunciante. Nesta fase de otimização deverá ser também analisado o site. O site deverá ser relacionado com os anúncios que o promovem, que por sua vez deverá ser relacionado com as palavras-chave. Só assim se garante que o tráfego recebido é qualificado e realmente interessado no produto ou serviço que o anunciante tem para oferecer.

2.3.4 Otimização ao nível de Palavras-chave

As palavras-chave deverão ser escolhidas de forma a serem o mais relacionadas possíveis com a campanha e a landing-page. Para maior visibilidade, deverão ser incluídas variantes relacionadas, versões singular e plural. Se necessário, poderão ser incluídos outros termos como métodos de escrita alternativas, sinónimos e erros ortográficos mais frequentes.

As palavras-chave poderão ser configuradas no seu tipo, ou *matching*. [ADRMTC]
Com algumas palavras será possível obter mais visualizações dos anúncios, enquanto

outras poderá ser obtido menos visualizações mas potencialmente mais cliques/tráfego qualificado. Assim, com uma boa estratégia de decisão de matching poderá ser aumentando a potencialidade da campanha, reduzindo até o CPC, e aumentando automaticamente o ROI.

<i>Tipo Matching</i>	<i>Exemplo</i>	<i>Descrição</i>
Broad	carros online	Frases similares e variações relevantes
Phrase	“carros online”	Frases similares que incluam a frase
Exact	[carros online]	Apenas e só a palavra-chave
Negative	-carros online	Que não incluam a palavra-chave (negação)

Tabela 3: Opções de Matching da Indústria

Neste momento, a plataforma de Anúncios do SAPO apenas permite utilizar palavras-chave cujo tipo de *Matching* é *Broad*, mas estudos recentes, confirmam que o *matching* das palavras-chave tem um papel crucial no retorno do investimento. [ONLmP]

2.3.5 Optimização ao nível de Anúncios

Os anúncios deverão ser o mais específico possíveis. Várias optimizações poderão ser feitas, novamente de acordo com o objectivo da campanha. No global, os anúncios deverão chamar a atenção do utilizador e apresentar, ao mesmo tempo, claramente o produto ou serviço do anunciante.

Para a optimização, é recomendado utilizar palavras nos anúncios, especialmente no título para apresentar aos utilizadores que o anúncio é relacionado com a sua pesquisa. Normalmente, se um enxerto do anúncio corresponder a uma pesquisa, o mesmo será apresentado a negrito destacando-o do anúncio e do resto da concorrência.

Como os utilizadores procuram mais por produtos específicos que marcas ou nome de empresas, deverá ser utilizado o espaço limitado a identificar claramente o produto em questão.

Uma vez que a cada pesquisa corresponde geralmente vários anunciantes concorrentes, é recomendado que se apresente a mais valia em relação à concorrência, descrevendo por exemplo características únicas ou promoções a oferecer.

A utilização de frases chave (ou *call-to-action*) deverá ser implementada em cada anúncio. Uma *call-to-action* encoraja os utilizadores a clicar no anúncio e certifica que o utilizador sabe exactamente com o que vai esperar ao chegar à página de destino. As

call-to-actions como “Compre Já”, “Visite Já”, “Registe-se” são normalmente as mais utilizadas.⁷

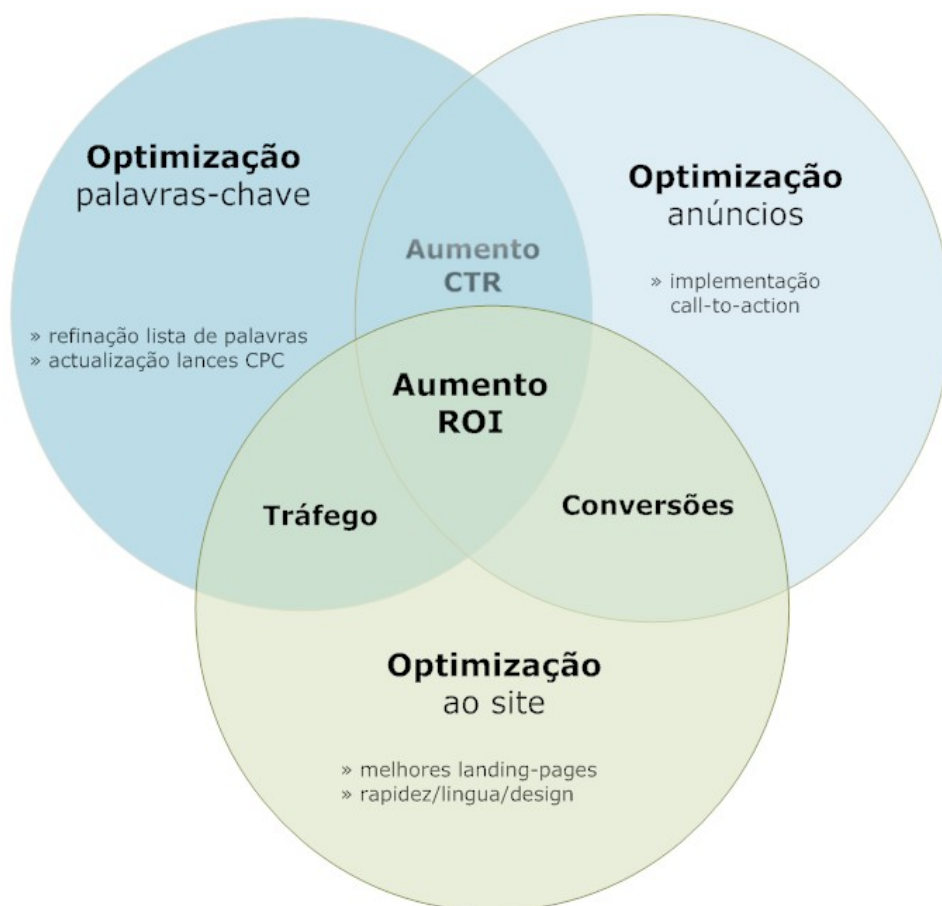
2.3.6 Optimização a Nível de Site

Independentemente de qualquer optimização, cabe ao site efectuar a conversão final. Assim, esta parte que em regra foge ao controlo de qualquer sistema de anúncios é a mais importante.

Para isso, no que respeita à conta de anúncios propriamente dita, deverá ser garantido que o anúncio possui como URL de destino a página do site correcta. É essencial que quando um utilizador clica num anúncio de texto, seja redireccionado à secção do site relacionada com o anúncio. Se os utilizadores não verificarem que a página do site seja relacionada com o anúncio, é muito provável que saiam do site sem qualquer conversão.

⁷ Existe claro situações que não requerem *call-to-actions*. Nos casos em que o anunciante apenas quer criar exposição à sua marca ou simples *branding* não se aplicam. Como são campanhas específicas e que em regra não se aplicam em motores de pesquisa, não serão analisadas.

Figura 4: Níveis de otimização de contas



A otimização do design do site, navegação, formulários e o conteúdo cabem exclusivamente ao anunciante.

2.4. Descrição do Problema e Trabalhos Relacionados

2.4.1 Introdução

Dada a crescente necessidade de rentabilizar as campanhas de anunciantes através de ferramentas próprias, as plataformas de anúncios em cenários de WEB necessitam de criar permanentemente novas métricas e algoritmos que possibilitam o aumento do ROI

aos seus anunciantes. Ao mesmo tempo, mantêm-se também actualizados face a um mercado de rápida mudança que é o da internet. Consegue-se assim fazer uma gestão mais eficaz e orientada ao rendimento do anunciante.

Para além de garantir uma gestão mais eficaz, a utilização de uma plataforma mais completa possibilita garantir mais segurança, prevenir fraude e permitir orientar o modelo de negócio optando por diferentes estratégias visto o seu panorama de performance da rede se encontrar muito mais completo.

2.4.2 Descrição do Problema

Na perspectiva de negócio, pretende-se tomar conhecimento de elementos individuais de uma determinada palavra-chave e de uma determinada campanha de forma a poder gerir e potencializar o retorno de investimento, nomeadamente:

- **Análise Comparativa** – Será relevante tanto para o anunciante como para a própria rede poder obter dados relativos que possibilitem comparar a performance de uma campanha ao nível das palavras-chaves utilizadas, em relação a outros anunciantes com temáticas semelhantes e à própria rede.
- **Análise Visual Gráfica** – A informação de performance de campanhas deverá ser apresentada aos anunciantes da maneira mais intuitiva e interactiva possível. É um facto que um anunciante normal lida, habitualmente, com uma grande quantidade de dados e informação que muitas vezes poderá ser considerada redundante ou irrelevante. Importa por isso identificar quais os dados mais relevantes a apresentar e de que forma. Interessará apresentar a informação o mais detalhada possível de acordo com o desejado.
- **Análise Interactiva** – Qualquer informação disponibilizada ao anunciante, para além de ser intuitiva deverá ser interactiva. Quer isto dizer que o anunciante deverá ter a possibilidade de interagir interrogando a Central de Anúncios de forma fácil e sempre que possível dentro dos próprios gráficos. A análise interactiva complementa a análise gráfica.
- **Optimização** – É crucial disponibilizar aos utilizadores sugestões de optimização. Assim, depois de definidos os objectivos de cada campanha pelo anunciante, deverão ser apresentados relatórios que o possibilitem tirar mais proveito da plataforma. As optimizações serão as referidas no ponto 5.5 e deverão ser calculadas por algoritmos definidos e em constante evolução.

2.4.3 Caracterização da Situação Actual do Anúncios SAPO

A situação actual da Central de Anúncios do SAPO, foi obtida durante a componente prática da dissertação, tendo sido feita uma avaliação a nível de Interface, Relatórios e Gestão de Negócio.

2.4.4 Interface

A Interface actual, acessível após a autenticação de um anunciante encontra-se por defeito em língua portuguesa estando possibilitado a alteração desta para o Inglês (Americano).

A Central de Anúncios está dividida entre duas áreas de navegação que são comuns a toda a actividade e uma área de conteúdo e contextualizada. A área de navegação encontra-se no topo que identifica o logotipo colaborativo, a navegação e o saldo. No fundo encontra-se localizado os links para os avisos legais, ajuda, termos de utilização e possibilidade de alterar a língua.

A área de conteúdo disponibiliza o *core* da Central de Anúncios. A gestão de campanhas, da conta, relatórios e outros é toda implementada nesta mesma área. Qualquer registo de opinião sobre o design propriamente dito e formatação de informação, seria meramente subjectivo, pelo que, para a escrita desta dissertação se mantém uma postura neutra, focando apenas na usabilidade.

2.4.5 Gestão de Conta

A Central de Anúncios do SAPO disponibiliza uma completa ferramenta da gestão de conta que permite, para além de editar os dados de conta propriamente ditos a qualquer momento, obter um historial completo de todas as actividades:

- **Histórico Detalhado (de Conta)** – Permite a consulta em formato de conta-corrente de toda a actividade contabilística da conta do anunciante. É possível verificar todos os créditos que são regra geral carregamentos de conta e débitos que corresponde às comissões auferidas pela prestação do serviço de anúncios. É possível definir um intervalo de dados para consulta.
- **Dados da Conta** – Dados pessoais e de facturação. O nome, morada, telefone, email e dados de contacto são definidos nesta zona. Como os dados de facturação são em muitos casos, especialmente nas empresas,

diferentes dos dados de contacto, o SAPO requer a introdução de dados diferentes.

- **Carregamentos (Pagamentos)** – Esta área de conta permite efectuar carregamentos em valor que se reflectem como um crédito no histórico detalhado aumentando o saldo/plafond disponível. Os pagamentos podem ser feitos por Multibanco, MBNET, Visa/Mastercard ou por Paypal. Poderá também ser definido uma periodicidade de pagamento, que de momento só pode ser única.

2.4.6 Gestão de Negócio

Um dos pontos mais importantes da Central de Anúncios do SAPO, é a gestão propriamente dita das suas campanhas. A interface desta secção deverá ser equilibrada tanto para um anunciante com pouco volume de palavras-chave/anúncios, como para um anunciante mais profissional, com uma solução mais completa. Assim, na Central de Anúncios, ao anunciante é dada a possibilidade de gerir as suas campanhas, grupos de anúncio, palavras-chave e anúncios.

O SAPO reconhece que esta secção deverá ser excepcionalmente rápida, intuitiva, e tem em mente que, em regra, os utilizadores terão a mesma metodologia na gestão das suas campanhas as quais se descreve:

- **Campanhas** – As campanhas correspondem ao primeiro nível de conta. Uma campanha pode ser um produto, um site/serviço. As campanhas vão agrupar grupos de anúncio que, por sua vez, agrupam anúncios específicos. Por exemplo, um anunciante de automóveis poderia criar uma campanha denominada “Vendas Automóveis”.
- **Grupo de Anúncio** – Estão no segundo nível de conta e directamente relacionados com as campanhas. Cada Campanha pode ter um ou mais grupos de anúncio. Voltando ao exemplo de um anunciante de automóveis poderia criar-se tantos grupos de anúncios como marcas de automóveis para a campanha “Vendas Automóveis”. Neste caso, cada grupo de anúncio era relacionado com a marca específica.

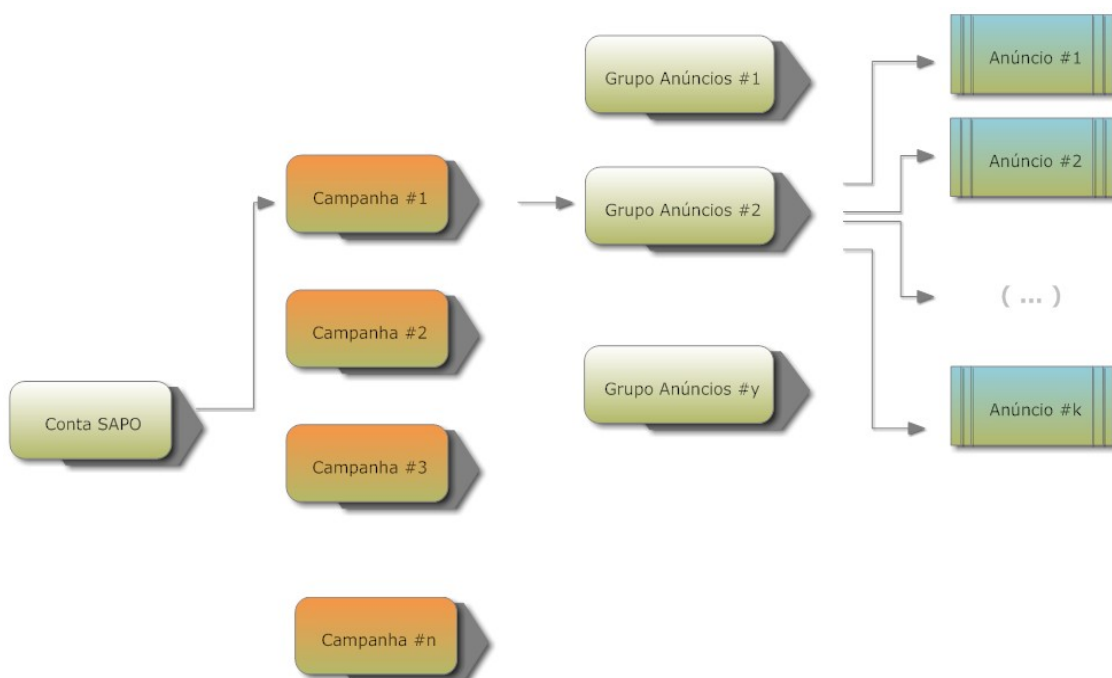


Figura 5: Relação Campanhas com Grupos de Anúncios com Anúncios

2.4.7 Relatórios

Os relatórios permitem aos anunciantes requisitar à plataforma dados de performance das suas campanhas. Para que os anunciantes tenham acesso a eles, deverão preencher um extenso formulário onde definem os dados a consultar, agrupamentos, períodos temporais e campanhas. Depois, deverão ser definidos os formatos em que o anunciante consulta os relatórios, nomeadamente em PDF, HTML, CSV ou HTML.

O anunciante dá um nome ao relatório e deverá aguardar que o mesmo esteja pronto. Este procedimento regra geral demora alguns minutos. Uma mais valia deste sistema de relatórios, é que o anunciante pode definir recorrências e gravar modelos. Isto é, o anunciante pode optar por semanalmente receber no seu email, um PDF com os relatórios da *performance* das suas campanhas nas últimas semanas. Para o projecto em questão, considera-se este modelo actualmente desactualizado e em certos casos obsoleto.

sapo.pt Internet SAPO Blogs Canais Classificados Fotos Mail Mapas Pesquisa Vídeos Notícias conta autenticada Logout

anúncios ANUNCIANTES

O seu saldo -0,-06 €

Campanhas Relatórios Conta Ferramentas Carregamentos Vouchers

Relatórios Novo Relatório

RELATÓRIOS

Relatórios Executados				
NOME DO RELATÓRIO	INTERVALO DE DATAS	DATA DO PEDIDO	ESTADO	AÇÃO
Novo Relatório				
gvsd	2011-05-03 - 2011-05-24	2011-05-24 09:26:03	read	Eliminar
fsdf	2011-05-09 - 2011-05-22	2011-05-23 10:33:00	error	Eliminar
nm	2011-04-26 - 2011-05-04	2011-05-04 16:26:19	read	Eliminar
Conta Abril	2011-04-01 - 2011-04-27	2011-04-27 17:36:10	read	Eliminar
adfit abril	2011-04-01 - 2011-04-27	2011-04-27 17:34:59	read	Eliminar
Mes a	2011-01-01 - 2011-01-17	2011-01-17 11:17:58	read	Eliminar
Mes	2011-01-01 - 2011-01-17	2011-01-17 11:01:56	read	Eliminar
a	2010-07-01 - 2010-07-07	2010-07-07 17:17:42		Eliminar
Mensal	2010-07-07 - 2012-07-13	2010-07-07 17:09:37		Eliminar
Mensal2	2010-05-01 - 2010-05-31	2010-07-07 17:08:45		Eliminar

|< << 1 2 >> >|

Modelos de Relatórios			
NOME DO RELATÓRIO	INTERVALO DE DATAS	PERIODICIDADE	AÇÕES
gvsd	2011-05-03 - 2011-05-24	active	Eliminar Executar
fsdf	2011-05-09 - 2011-05-22	active	Eliminar Executar
nm	2011-04-26 - 2011-05-04	active	Eliminar Executar
Conta Abril	2011-04-01 - 2011-04-27	active	Eliminar Executar
adfit abril	2011-04-01 - 2011-04-27	active	Eliminar Executar
Mes a	2011-01-01 - 2011-01-17	active	Eliminar Executar
Mes	2011-01-01 - 2011-01-17	active	Eliminar Executar
a	2010-07-01 - 2010-07-07	active	Eliminar Executar
Mensal	2010-07-07 - 2012-07-13	active	Eliminar Executar

Figura 6: Relatórios do Anuncios SAPO

2.4.8 Conclusões

A Plataforma de Anúncios do SAPO possui de momento o essencial para criar e disponibilizar campanhas. Existem ferramentas bem intuitivas para criar campanhas, grupos de anúncio e texto. As restrições e limitações estão apresentadas de forma simples, e sempre que possível utiliza-se o AJAX para tornar os acessos, nomeadamente as tarefas mais repetitivas como gerir anúncios, mais rápidas e eficazes.

Contudo, no que diz respeito às optimizações e acompanhamento de *performance* há lacunas por resolver. Prudentemente, o SAPO optou por incorporar um importador de campanhas do Google Adwords, pois em regra, os anunciantes optimizam as suas contas nesse sistema para mais tarde importar.

Como referências de concorrência a actuar no mercado português, destaca-se o Google Adwords⁸, o Yahoo Search Marketing⁹ e o Microsoft Adcenter¹⁰ (muito embora actualmente tenha sido substituído na rede Bing[BING] pelo anúncios SAPO. [BINGSAPO]).

8 Google Adwords em <http://adwords.google.com> [30/05/11]

9 Yahoo Search Marketing em <http://searchmarketing.yahoo.com/> [30/05/11]

10 Microsoft Adcenter em <https://adcenter.microsoft.com> [30/05/11]

3. Trabalhos Relacionados

3.1. Modelos de Visualização

Os modelos de visualização permitem de forma visual e gráfica, oferecer aos utilizadores uma visão global ou específica dos seus dados, sejam de uma dimensão, ou de n-dimensões. O utilizador ao ter os seus relatórios de forma não-estática, interactiva e visualmente mais agradável, consegue ter uma percepção mais exacta e actualizada, detalhando o relatório o quanto deseja. Aqui apresentamos três soluções que poderão ser implementadas com o objectivo de tornar a informação textual em gráfica.

3.1.1 Polaris

O Polaris, é uma framework de visualização gráfica com o objectivo de apresentar tabelas multidimensional de grandes bases de dados. É específico para bases de dados e possui uma interface do estilo "drag-and-drop" onde as tabelas são arrastadas para o modelo gráfico. No Polaris, é estimulado o processo de interactividade para uma melhor leitura de gráficos complexos. Este modelo foi desenvolvido com o propósito de apresentar ao utilizador um mecanismo gráfico para obter leituras, como tendências, padrões ou sazonalidade. Ideal para investigação, com um grande formalismo matemático, desejado em grandes análises.

3.1.2 Infovis

A Toolkit Infovis foi apresentada no paper de Jean-Daniel Fekete (universidade de Paris-Sud), desenhado para suportar criação e integração de modelos a duas dimensões recorrendo a Java.[INFOVS] As principais características desta Toolkit, corresponde a uma óptima utilização de memória, traduzindo-se na possibilidade de análise de um grande número de dados. Para além disso, é bastante rápido, pois recorre ao sistema de aceleração gráfica Agile2D[AGILE2d], que corresponde a uma implementação do Java2D baseado em Open GL API para aceleração gráfica por hardware.

3.1.3 Protovis

Já Michael Bostock e Jeffrey Heer, identificam no artigo "Protovis: A Graphical Toolkit for Visualization"[PROMBJH], o desequilíbrio entre flexibilidade na apreciação gráfica e dificuldade na implementação, contra uma certa inflexibilidade e implementação mais fácil. Assim, soluciona-se alguns dos problemas quando se criam módulos independentes: o tratamento dos dados em análise, e a visualização gráfica. Ao criar mecanismos independentes, obtém-se toda a liberdade de configuração e tratamento gráfico, mantendo uma estrutura modular, integrando-se ambos por standards. O Protovis é apresentado como uma resposta a este paradigma, ao ser criado numa linguagem de configuração própria, orientada apenas ao grafismo, deixando todo o tratamento de dados independente. Agora, existe a possibilidade de manter, por exemplo, uma equipa dedicada ao tratamento de dados, e outra dedicada apenas ao grafismo. Com o protovis os programadores/integradores obtêm uma grande flexibilidade, em particular, pelo uso de marks, anchors, panels e data transformations. Funcionando em Javascript e com uma integração possível de ser feita em JSON, o protovis, sendo baseado em modelos CSS é uma ferramenta que poderá ser bastante útil na criação de visualização gráfica para ambientes WEB, mesmo mantendo uma interacção gráfica bastante presente com n-variáveis.[WWWPROT]

3.1.4 Conclusão

Em conclusão deste capítulo, em termos objectivos e independentemente das suas aplicações neste projecto, poderemos comparar as três ferramentas aqui referenciadas:

<i>Ferramenta</i>	<i>Última Versão</i>	<i>Plataforma</i>	<i>Tipo Licença</i>
Protovis ¹¹	3.2	Javascript+SGV	BSD License (open source)
Infovis ¹²	2.0.1	Javascript+SGV	Open Source
Polaris ¹³	2.0	Windows	Comercial

3.2. Ferramentas de Auxílio de Negócio e Optimização

As ferramentas de auxílio de negócio, são constituídas por acrescentos à Central de Anúncios, não essenciais para o seu funcionamento básico, mas que permitem ao anunciante de certa forma facilitar ou automatizar o seu trabalho.

Essas ferramentas são uma mais-valia a qualquer rede de afiliação pois ao partir de uma premissa em que o tempo de qualquer anunciante é precioso, e que o mesmo é melhor investido na optimização de campanhas e não em tarefas metódicas, podemos assim garantir que enquanto existe uma interacção do Anunciante com a Central de Anúncios, a mesma é investida na optimização de campanhas propriamente dita.

Por outro lado, qualquer mecanismo que auxilie o anunciante de forma (às vezes) transparente a optimizar as suas campanhas poderá vir a ser implementado. Sistemas esses, que utilizando algoritmos para o efeito, refinem as listas de palavras-chave, propõem alterações a CPCs e até anúncios. As ferramentas de optimização deverão ser orientadas ao ROI e cabe ao anunciante a decisão final de implementar ou não as sugestões dadas.

3.2.1 Sugestão de Palavras Chave

Uma das funções dos anunciantes numa central de anúncios, é a gestão de palavras-chave. Importa definir todas as palavras-chave possíveis relacionadas com o anúncio. Se estivermos a gerir uma campanha de venda de automóveis, importa manter uma lista de palavras-chave com termos relacionados com automóveis, como “automóvel”, “carros”, “veículos”, etc., mas não “carros telecomandados”, ou “carros de colecção” ou outras palavras-chave que desvirtuem o sentido do anúncio, ou que embora estejam relacionados com o tema, não definam o produto ou serviço à venda. Assim, é imperativo implementar um sistema de sugestões de palavra-chave, que auxilie o anunciante na

¹¹ Protovis – A Graphical Approach to Visualization - <http://vis.stanford.edu/protovis/>

¹² JavaScript InfoVis Toolkit – Create Interactive Data Visualizations for the Web - <http://thejit.org/>

¹³ Polaris: Database and Data Cube Visualization - <http://graphics.stanford.edu/projects/polaris/>

constante busca e refinação das suas listas. Esta ferramenta será essencial sobretudo em anunciantes estrangeiros que dominem mal o mercado em que estão a operar, ou em tópicos que não sejam do conhecimento do anunciante. É fácil assumir que automóvel é sinónimo de carro, mas poderá não ser tão claro para um anunciante estrangeiro.

No paper de Kevin Bartz e Cory Barr, “Natural Language Generation for Sponsored-Search Advertisements” [NLSSA] são identificados os possíveis algoritmos a aplicar, assim como os obstáculos. Trabalhar em linguagem natural pode ser complicado, especialmente numa língua como o português onde uma simples alteração de uma vírgula, poderá alterar todo o sentido de uma frase, às vezes até ao oposto.

São assim identificados três passos a trabalhar, que se destacam:

Passo	Dados Introduzidos	Dados Correctos	Dados Errados
Ordenar, inserir palavras	Fénix ordem	Ordem de fénix	Ordens fénix
Adicionar modificadores	Ordem de fénix	Ordem de fénix, livros	Ordem de fénix
Capitalizar	Ordem de fénix, livros	Ordem de Fénix, livros.	Ordem de Fénix, Livros

Tabela 4: Passos propostos do algoritmo

Os autores propõem que os dados de “treino” para o refinamento e sugestão de palavras chave sejam obtidos por pesquisas (foram utilizados 12 mil milhões de pesquisas reais), e para a capitalização, seja utilizado o algoritmo de M Porter [MPORTES]

3.2.2 Optimização de Orçamento Publicitário (*Budget*)

Um grande problema de qualquer anunciante, é como pode maximizar o seu orçamento que tem pré-definido para as suas campanhas. Depois de um anunciante definir um orçamento, seja ele de que valor for, deverá ser gasto na maneira vantajosa para os seus objectivos. Assim, poderão ser aplicados algoritmos que auxiliem na gestão do *budget*.

3.2.3 Estimativa de CTR com regras de decisão

Relativamente às estimativas de CTR¹⁴, existem dois grande problemas que os anunciantes enfrentam. O primeiro problema consiste em estimar o CTR de uma campanha baseando-se em dados históricos, sendo o segundo, estimar o CTR de novos anúncios, ou anúncios com pouco volume de tráfego.

- **Estimação de CTR** – O grande problema em comparar anúncios, é que gostaríamos de utilizar um conteúdo do anúncio. Mas infelizmente, os anúncios são apresentados em diferentes posições de uma página e em diferentes contextos o que, por si só, cria uma acentuada alteração ao CTR. Assim, será proposto estimar pelo “Maximum Likelihood Estimation (MLE)” para apresentar a estimativa de CTR mediante dados históricos.
- **Previsão de CTR** – Para prever um CTR em anúncios novos, ou com históricos escassos, deveremos utilizar uma outra estratégia. Assim, iremos aplicar o algoritmo de RuleFit, SLIPPER e LRI por exemplo.

No paper “Predicting Ads’ ClickThrough Rate with Decision Rules” criado por três docentes da faculdade de Poznan – Polónia,[PREDICTCTR] são apresentadas soluções de implementação e de estudo para obter uma estimativa CTR.

3.2.4 Optimização a Nível do Website do anunciante

A optimização a nível da conversion-rate é uma arte ou até uma ciência que consiste em persuadir os visitantes dos nossos sites em ter acções do nosso interesse, sejam elas a compra de algum produto, leitura de um PDF, donativo ou mesmo um simples contacto.

As optimizações a nível do site, são inúmeras e têm como objectivo aumentar as vendas, os lucros, ou mesmo acções específicas no site. Existem inúmeras optimizações a serem efectuadas, e estas saem um pouco dos objectivos de uma central de anúncios, mas mesmo assim, deverá ser apresentada de forma informal, sugestões que se considerem críticas para o bom sucesso.

14 CTR – Click-Through Rate – Taxa de cliques por 100 visualizações. A um CTR de 5%, deduzimos que em cada 100 visualizações de um determinado anúncio, obtemos 5 cliques no mesmo.

No livro “Website Optimization” publicado pela O'Reilly pelo autor Andre B.King[WEBSITEOPT] são apresentadas inúmeras sugestões a ter em conta na optimização dos sites:

- **Web Page Optimizion** – Esta optimização será mais técnica e orientada para os factores físicos do site: velocidade, performance e estruturação. Os sites deverão abrir sempre bastante rápidos, não possuir código redundante, e sempre que possível, os servidores deverão estar fisicamente localizados na mesma zona geográfica que o público alvo. - [Página 155 a 185]
- **CSS Optimization** – Uma boa estrutura de CSS, providencia uma abertura mais rápida do site. Existem muitas técnicas de CSS que bem aplicadas demonstram tornar um site muito mais rápida, graças à correcta utilização das hierarquias de CSS [Página 186 a 214]
- **Ajax Optimization** – A nova metodologia de AJAX, torna os sites mais interactivos e agradáveis, mas também os podem tornar mais eficazes e rápidos. Regra geral, sempre que existe um pedido de informação por parte dos visitantes, o servidor e o browser deverão apresentar a página WEB completa. Graças ao AJAX, apenas a informação numa posição específica é que será apresentada. Assim, ao evitar que os sites sejam sempre carregados a qualquer interacção, temos uma experiencia de navegação bastante melhor [Página 217 a 256]

4. Desenvolvimento do Projecto

4.1. Estrutura do Documento

Neste capítulo, o estado de arte dos principais conceitos e tecnologias a ser utilizados serão definidos, nomeadamente:

- Pré-processamento e filtragem de dados brutos de dados relativos a campanhas
- Relatórios Visuais que se pretendem implementar sobre a actual Central de Anúncios do SAPO

4.2. Pré-Processamento e Tratamento de Dados

Os dados utilizados no projecto, são dados reais obtidos directamente dos servidores de anúncios do SAPO. Os dados foram comunicados por canais seguros, utilizando encriptação PGP.

Os dados fornecidos, correspondem a cerca de 16 milhões de registos que correspondem à totalidade do volume do periodo de 1 de Novembro de 2010 a 31 de Dezembro desse mesmo ano. Esses dados, estão distribuidos entre várias Tabelas SQL que correspondem à estrutura real em pleno uso pelo SAPO. É de salientar o facto de muitos desses dados estarem já agrupados em certas tabelas. Quer isto dizer que para

consultar o somatório de impressões ou de lances de uma determinada palavra-chave e de uma determinada campanha, basta ser consultada a tabela respectiva.¹⁵

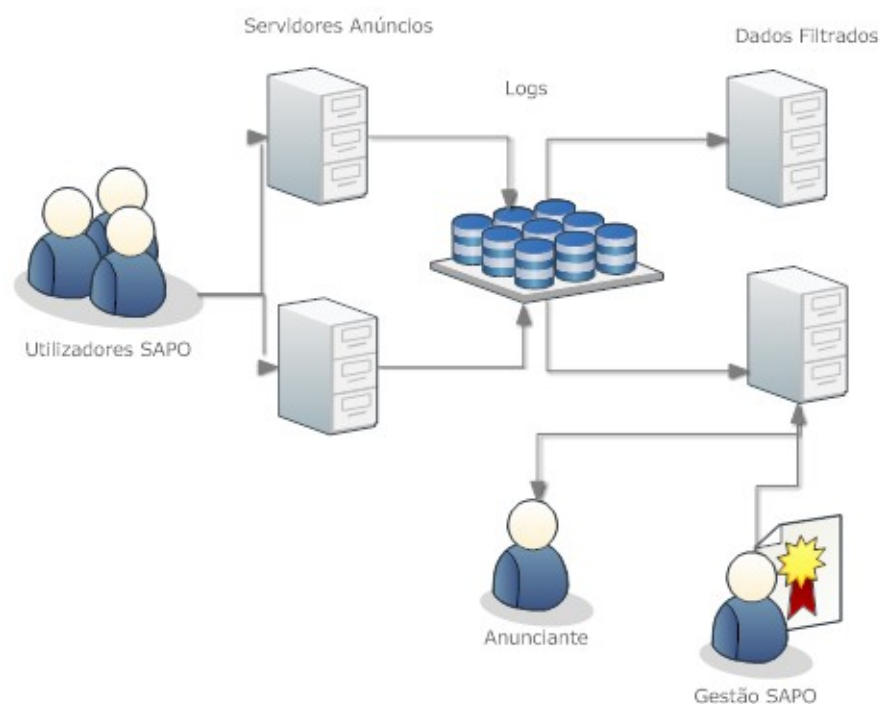


Figura 7: Estrutura e Relação Dados Brutos/Filtrados

¹⁵ Neste caso, podemos consultar a tabela wrd_performance_report que nos fornece dados pré-agrupados por palavra.

4.2.1 Tipo de Dados

Os dados actuais também utilizados para o desenvolvimento do projecto, são no formato SQL numa base de dados MySQL com o motor InnoDB.

	Table	Action	Records	Type	Collation	Size
A	adgroup_word		0	InnoDB	utf8_general_ci	64.0 KiB
	adg_adgroups		0	InnoDB	utf8_general_ci	48.0 KiB
	Adsense_AffiliateSite		0	InnoDB	utf8_general_ci	48.0 KiB
	Adsense_PositionInSite		0	InnoDB	utf8_general_ci	80.0 KiB
	ads_ads		0	InnoDB	utf8_general_ci	48.0 KiB
	bid_impressions		~5,568,516	InnoDB	utf8_general_ci	615.1 MiB
	cam_campaigns		0	InnoDB	utf8_general_ci	48.0 KiB
	clicks		~2,109,629	InnoDB	utf8_general_ci	259.9 MiB
	conversions		~31,151	InnoDB	utf8_general_ci	5.5 MiB
	site_impressions		~24,697	InnoDB	utf8_general_ci	3.0 MiB
B	usr_impressions_report		~8,098,786	InnoDB	utf8_general_ci	2.1 GiB
	wrd_adg_performance_report		~247,223	InnoDB	latin1_swedish_ci	21.6 MiB
	wrd_performance_report		~112,493	InnoDB	latin1_swedish_ci	5.5 MiB
13 table(s)		Sum	~16,192,495	MyISAM	latin1_swedish_ci	2.9 GiB

A - Tabelas SQL de dados disponibilizadas

B - Tabelas SQL implementadas

Figura 8: Resumo de tabelas SQL

Os dados fornecidos consistem em cerca de 1,583,779 registos, correspondem a logs dos servidores de anúncios SAPO entre o periodo de 1 de Novembro de 2010 e 31 de Dezembro de 2010. Na Figura 8, apresentam-se os dados agrupados com inserções redundantes.

Para além destes dados já tratados, será também implementado uma arquitectura paralela que dará mais flexibilidade à realização de relatórios visuais. Por exemplo, o actual sistema impossibilita o cruzamento ou segmentação de dados geográficos.

É impossível verificar se houve ou não um clique, conversão ou pageview na região do Porto.

4.2.2 Critérios de Refinação

Pela análise de mercado, sugestões de anunciantes e padrões da indústria, identificaram-se uma série de dados que deverão ser “filtrados” e tratados do seu estado bruto. Esses dados existem e sempre existirão no futuro, mas não estão implementados nos relatórios textuais do momento.

- Localização geográfica do autor (coordenadas Longitude/Latitude)

- Dados técnicos como sistema operativo, linguas, versão do Javascript entre outros
- Tendências do utilizador, analisando o seu historial passado. Quando visitou e o quê no passado
- Fidelidade ao analisar quantas vezes o utilizador visitou ou que outros temas relacionados já foram pesquisados pelo utilizador.

Importa assim definir o tratamento desses dados e a melhor forma de os cruzar ou segmentar em informações contextualizadas que permitam ao anunciante as melhores decisões de negócio. Assim, introduzimos o conceito de “Relevância” onde procuraremos obter a melhor informação de acordo com o seguinte diagrama de *Venn*:



Figura 9: Diagrama de Venn para a Relevância

Significa assim que a Relevância é a intersecção da informação do “Termo de Pesquisa”, com a “Informação Geográfica” e com a “Sessão/Fidelidade”.

4.3. Relatórios Visuais

Esta secção disponibiliza quais os factores a evoluir, em tudo o que se refere à maneira como a informação chega ao anunciante e quais os métodos que este dispõe para interagir nos seus dados.

4.3.1 Informação Disponibilizada

Actualmente os utilizadores do *back-office* do Anúncios SAPO têm acesso a informação que lhes é transmitida de forma textual e descontextualizada. A informação das suas campanhas é textual porque, em caso algum, são fornecido gráficos a acompanhar os dados, e ao mesmo tempo é descontextualizada porque não existe possibilidade de os cruzar ou segmentar. Por exemplo, não é possível obter um relatório de dados entre datas específicas.

4.3.2 Integração

A informação a disponibilizar será interactiva e actual, utilizando para isso várias APIs de Javascript que irão garantir que os dados disponibilizados são apresentados de forma gráfica e intuitiva.

Assim, de uma actual interface meramente textual e com informação difícil de obter leituras:

The screenshot displays the SAPO Anúncios user interface. At the top, there is a navigation bar with links to Internet SAPO, Blogs, Canais, Classificados, Fotos, Mail, Mapas, Pesquisa, Vídeos, and Notícias. A user is logged in, as indicated by 'conta autenticada' and a 'Logout' button. The main header shows the 'anúncios' logo and a balance of 395,92 €. Below this, there are tabs for 'Campanhas', 'Relatórios', 'Conta', 'Ferramentas', 'Carregamentos', and 'Vouchers'. The 'Campanhas' tab is active, showing sub-tabs for 'Campanhas Activas', 'Todas as campanhas', 'Campanhas Suspensas', and 'Nova Campanha'. The 'Todas as Campanhas' section contains a table of active campaigns.

NOME DA CAMPANHA	ESTADO	SEGMENTAÇÃO	ORÇAMENTO DIÁRIO	CLIQUEs	IMPRESSÕES	CONVERSÕES	CPC MÉDIO	AÇÃO
Nova Campanha								
CN - Adelgafit	Activa	Site	25.00 €	2911	2371821	37	0.20 €	Suspender
CN - Espionagem.com	Activa	Site	25.00 €	516	1229760	2	0.20 €	Suspender
KW - Espionagem.com	Activa	Palavras-chave	25.00 €	652	252651	4	0.08 €	Suspender
KW - Insania	Activa	Palavras-chave	25.00 €	555	1095	0	0.08 €	Suspender
[PT] - Adelgafit	Activa	Palavras-chave	100.00 €	8931	6225988	127	0.36 €	Suspender
TOTALIS				20.395	35.565.229	187		

At the bottom, there is a footer with links for 'Ajuda / Contactos', 'Termos de Utilização', and 'Blog oficial'. It also includes a small logo and text stating 'Produzido por PTC © 2010. Todos os direitos reservados. O SAPO é uma marca e um motor de busca criado na Universidade de Aveiro.'

Figura 10: Interface Actual do Anúncios SAPO

Pretende-se implementar uma interface mais evoluída e acompanhada de gráficos interactivos:

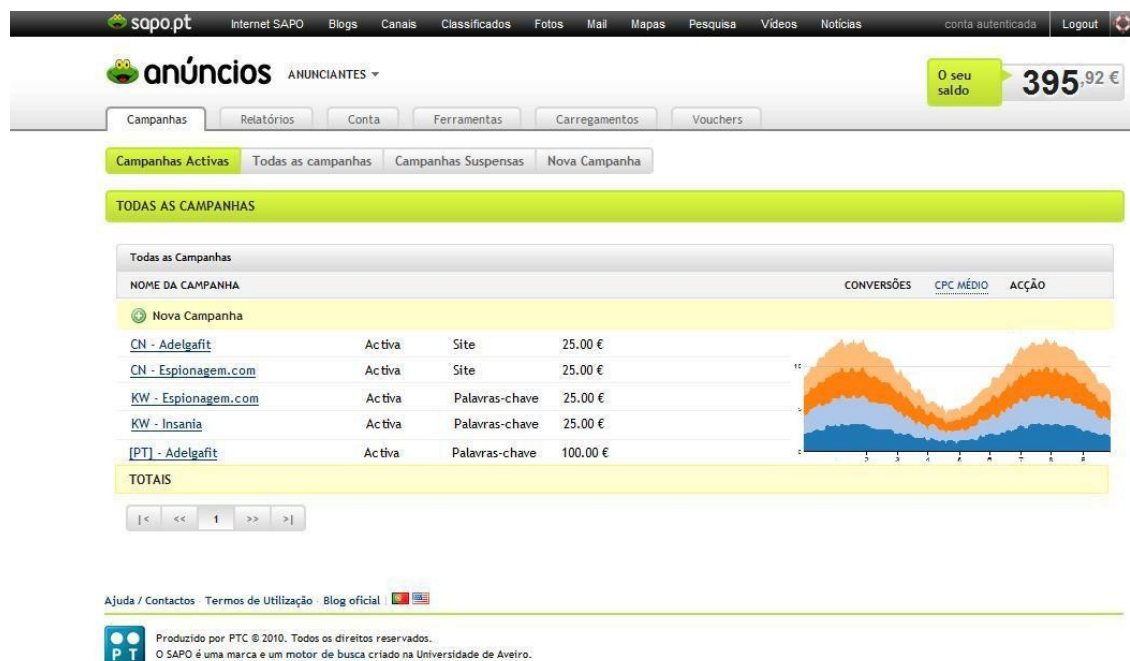


Figura 11: Interface Desejada do Anúncios SAPO

4.3.3 Soluções Adoptadas

Para a implementação dos relatórios visuais, sistemas de pré-processamento e de optimização, foram implementadas as tecnologias PHP e Javascript com a utilização das frameworks JQUERY para todo o código desenvolvido em Javascript, e PROTOVIS para a correcta implementação de gráficos interactivos. Foram estudadas soluções já adoptadas pelo próprio SAPO, mas optou-se pela utilização de outras ferramentas com outras características as quais são aqui apontadas. Como outras possibilidades de implementação, foi abordado o Axiis[AXIIS] que apresentou uma versão BETA, cuja última versão foi em 2010. O facto da última versão ser de 201, poder ter como significado que não deverá ser futuramente evoluída. Além disso, os seus gráficos são

gerados em Flash, e por ser inteiramente dependente de *plugins* ao browser se optou por descartar.

Assim, iremos implementar o PROTOVIS[PROMBJH] que consiste numa *toolkit* de visualização para Javascript que possibilita a visualização gráfica.

5. Implementação e Resultados Obtidos

Para a implementação do projecto, foram utilizados diversos servidores dedicados, nomeadamente um servidor MySQL com o sistema de tabelas InnoDB, o Apache a ser executado com o PHP 5.3 e o phpMyAdmin para a configuração das tabelas.

5.1. Alterações às Base de Dados

As tabelas SQL que foram fornecidas, necessitaram de algumas alterações. Assim, foram em certos casos feitas divisões de campos, e noutros adicionados outros campos.

Na tabela “*conversions*”, uma vez que o campo `conv_time` guardava data e hora no mesmo campo (DATETIME), optou-se por criar um novo campo, chamado `conv_date` e registar em `conv_time` as horas e no `conv_date` a data da conversão:

```
ALTER TABLE `conversions` ADD `conv_time` TIME NOT NULL AFTER  
`conv_date` ;ALTER TABLE `conversions` DROP PRIMARY KEY ,  
  
ADD PRIMARY KEY ( `conv_date` , `bid_id` , `ad_id` , `conv_time` );  
UPDATE `conversions` SET conv_time = time( conv_date ) ;ALTER TABLE  
`conversions` CHANGE `conv_date` `conv_date` DATE NOT NULL
```

Foram criadas as tabelas wrd_adg_performance que regista por palavra-chave e por adgroup, as métricas, individualizadas ao adgroup, correspondendo ao Custo total, conversões totais, impression share, click-share, conversion-share e custo share.

```
CREATE TABLE IF NOT EXISTS `wrd_adg_performance_report` (  
  
  `wrd_id` int(11) NOT NULL,  
  
  `adg_id` int(11) NOT NULL,  
  
  `impressoes` int(10) unsigned NOT NULL default '0',  
  
  `clicks` int(10) unsigned NOT NULL default '0',  
  
  `ctr` float(7,2) NOT NULL default '0.00',  
  
  `conversions` int(10) unsigned NOT NULL default '0',  
  
  `custo` float(7,2) NOT NULL default '0.00',  
  
  `avg_posicao_media` float(7,2) NOT NULL default '0.00',  
  
  `imp_share` float(7,4) NOT NULL,  
  
  `click_share` float(7,4) NOT NULL,  
  
  `conv_share` float(7,4) NOT NULL,  
  
  `custo_share` float(7,4) default NULL,  
  
  `pontuacao` int(11) NOT NULL,  
  
  PRIMARY KEY (`wrd_id`,`adg_id`)  
  
  ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
```

5.2. Sistemas de *pre-parsing*

Desenvolvida a estrutura SQL que permite suportar todos os dados, foi criando um conjunto de scripts que efectuam a leitura de todos os dados, os agrupam, calculam e depois inserem esses mesmos resultados nas novas tabelas. São esses dados que serão lidos, de forma praticamente instantânea, pois as tabelas estão optimizadas para leitura,

que serão acedidas pelos algoritmos desenvolvidos que implementam as métricas e devolvem os dados a serem consultados nos relatórios. Esses scripts deverão ser executados remotamente, de forma a manter os dados frescos num periodo constante, por exemplo, a cada 20 minutos ou se o fluxo de rede se justificar, a cada minuto.

5.3. Visualização de Relatórios

Passamos agora para a apresentação dos gráficos desenvolvidos, propriamente ditos. São 3 fontes de dados, e cada gráfico possui um objectivo na leitura dos relatórios:

5.3.1 Gráfico Impressões + Cliques + Conversões

Este gráfico, apresenta entre dois limites temporais a evolução multidimensional das impressões, cliques e conversões efectuadas. Pode também ser apresentado o custo respectivo e podem ser dadas restrições como o Site ID, posição e grupo de anúncio.

Centro de Visualização de Anúncios SAPO

Relatórios Optimizador



Figura 12: Gráfico Linear Impressões, Cliques e Conversões (versão normal)

Pode também ser especificado um tipo de segmentação que poderá ser por dia, por semana ou dia da semana. Esta segmentação é útil para identificar, por exemplo, padrões temporais e sazonalidade na performance das campanhas.

Centro de Visualização de Anúncios SAPO

Relatórios Optimizador

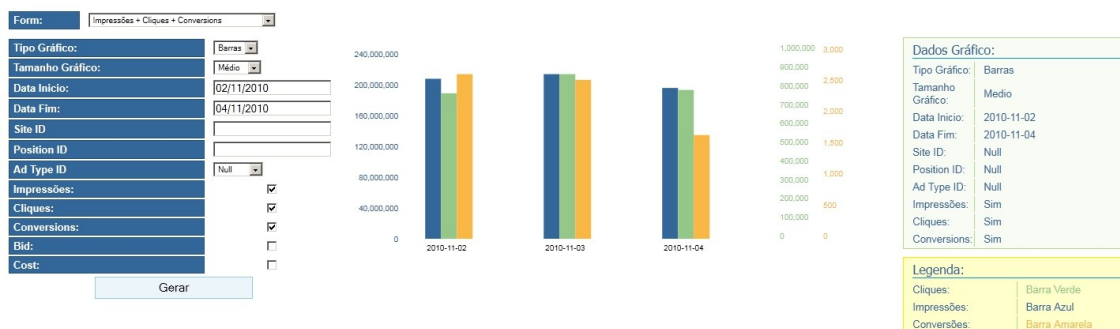


Figura 13: Gráfico Barras de Impressões, Cliques e Conversions

Centro de Visualização de Anúncios SAPO

Relatórios Optimizador



Figura 14: Gráfico Linear com Condicionantes

Também existe a possibilidade de gerar gráficos em forma de barras, tal como apresentado na Figura 13

Centro de Visualização de Anúncios SAPO

Relatórios Optimizador

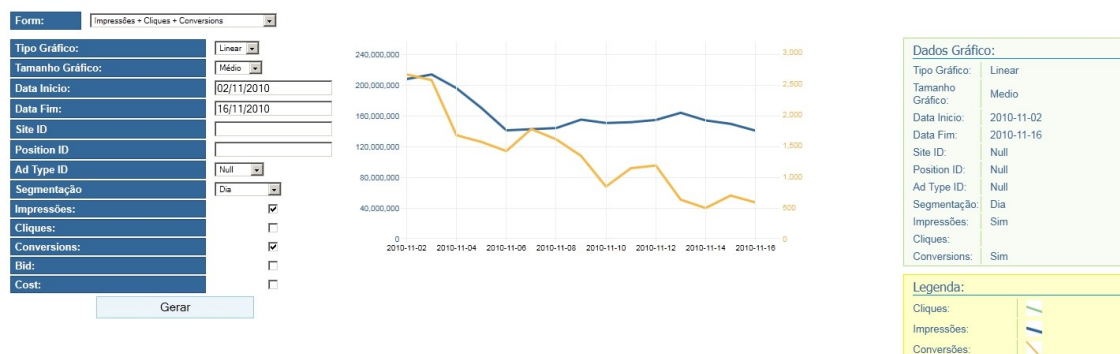


Figura 15: Gráfico Linear com Condicionantes 02/11 a 16/11

Podemos aplicar condicionantes com outras datas, tal como apresentado na Figura 15.

5.3.2 Gráfico CTR com Conversion Rate e Impression Share



Figura 16: Gráfico de Conversion Rate com CTR e Impression Share (versão normal)

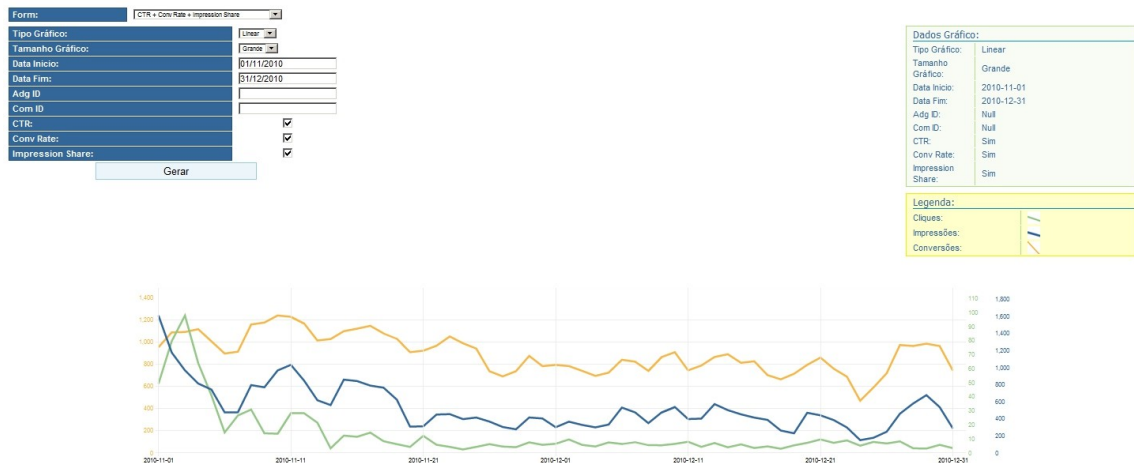


Figura 17: Gráfico de Conversion Rate com CTR e Impression Share (versão maior)

5.3.3 Gráfico Scatterplots Adg ID + Wrd ID

Os gráficos Scatterplots, ou gráficos de dispersão, são representações em duas ou mais variáveis, cada uma em função da outra. Graças a este gráfico, é possível o anunciante ter uma visão global e interligada das performances das suas estatísticas, nomeadamente como está por exemplo o CTR em relação ao seu CPC. É um gráfico útil, para novas palavras-chave ou para ver a performance individualizada de palavras-chave ou grupos de anúncios.

Neste caso, podemos relacionar numa só vista para um determinado adgroup, as keywords. Onde o diâmetro dos círculos e a localização x/y dos mesmos, demonstram a relação entre visualizações, cliques e conversões.

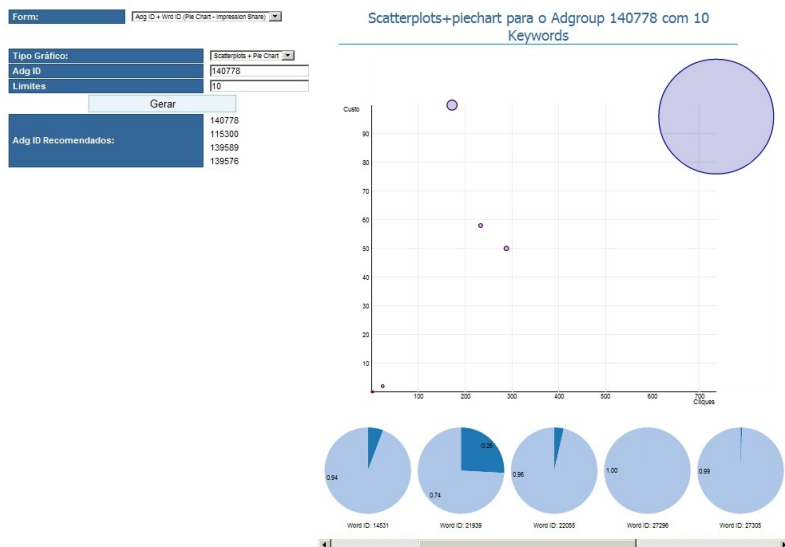


Figura 19: Scatterplots + Piechart com 10 keywords

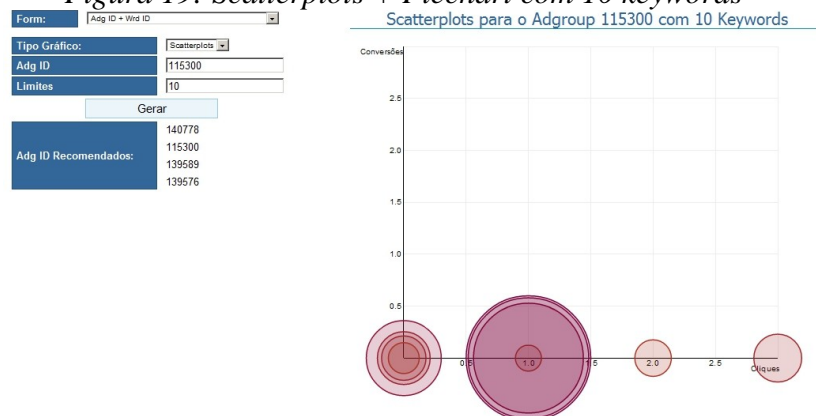


Figura 18: Scatter plots com 10 keywords

Agora, na presente figura, demonstramos como podemos relacionar as keywords com os CTR, impressões e cliques. Quanto maior o diâmetro do círculo, maior CTR apresenta. No fundo, voltamos a disponibilizar as mesmas palavras-chave, numa *pie-chart* onde apresentamos a relação de Impression-Share. Por exemplo, uma palavra-chave com um CTR alto e uma impression-share alta, como o caso da apresentada na figura, demonstra uma boa gestão da mesma.

5.3.4 Optimizador

O optimizador implementado baseia-se no algoritmo que compara os dados de uma determinada campanha, e no periodo de 30 dias (periodo esse que poderá ser alterado, caso a caso), com a performance da rede, no conjunto de palavras chave. Assim, comparando os termos CTR, valor CPC e as métricas de *share* apresentadas em

Centro de Visualização de Anúncios SAPO

[Relatórios](#)[Optimizador](#)

Optimizador de Campanhas Anúncios SAPO

Para a palavra-chave **254530** e adgroup **139576** para os últimos 30 dias

Optimização

Média CTR para a KW	0.10%
Média CPC Custo	€ 0,05
Posição Média	1.84
Conversões	€ 0,00/conversão (0%)
CTR do Adgroup	3.91%
CPC do Adgroup	€ 0,05
Impression Share	0%
Conversões	€ 0,07/conversão (70.43%)
Posição Média	1
Sugestão:	
<ul style="list-style-type: none">• A sua "posição média" está ok.• A sua "conversion rate" encontra-se adequada.• Outros anunciantes licitam com maior CPC.• CTR superior à média	

capítulos anteriores, sugere o tipo de optimização a fazer.

Este processo é apenas sugestivo/informativo, isto é, qualquer alteração a fazer terá que ser pelo próprio anunciante sendo a sua gestão autónoma. Neste contexto, o sistema de optimizações, a uma dada palavra chave e um dado adgroup, apresenta amostras reais, entre os dados da palavra-chave específica, e a média da rede, nomeadamente a média CTR para a palavra-chave, o CPC médio pago por essa palavra-chave, a sua posição média e a conversion rate.

Paralelamente, são apresentados dados da rede para as mesmas palavras-chave de terceiros (sem obviamente nunca identificar um anunciante em específico).

No final, é feito um relatório textual onde informa verbalmente o estado da posição média, conversion share, CTR e valores de custo. Optou-se por ser textual, pois nesta fase, pretendemos ser o mais transparentes possíveis ao anunciante evitando termos técnicos.

5.4. Análise de Resultados

Com a implementação do projecto em dados reais, foi possível efectuar uma análise crítica do seu resultado. Foram identificadas lacunas que por falta de possibilidade de alteração ou a nível tecnológico (recorde-se que o SAPO foi bastante restrito nos dados e condições oferecidas para a realização do projecto) não foram implementadas.

5.5. Sugestões Automáticas de Optimização

Após o anunciante ter informado os seus objectivos de ROI, target CPA ou outros equivalentes, é possível a um sistema estudar e analisar os dados existentes e propor optimizações que correspondam a acertos de CPC, CPA ou outros, assim como optimizações das páginas de destino.

5.6. Redefinição de Navegação

Com a implementação de novas ferramentas e metodologias de visualização, será prudente uma actualização à navegação actual. Assim, considerando que a actual navegação é adequada para a situação actual, vamos apenas actualizar a mesma à nova realidade aqui proposta.

5.6.1 Campanhas

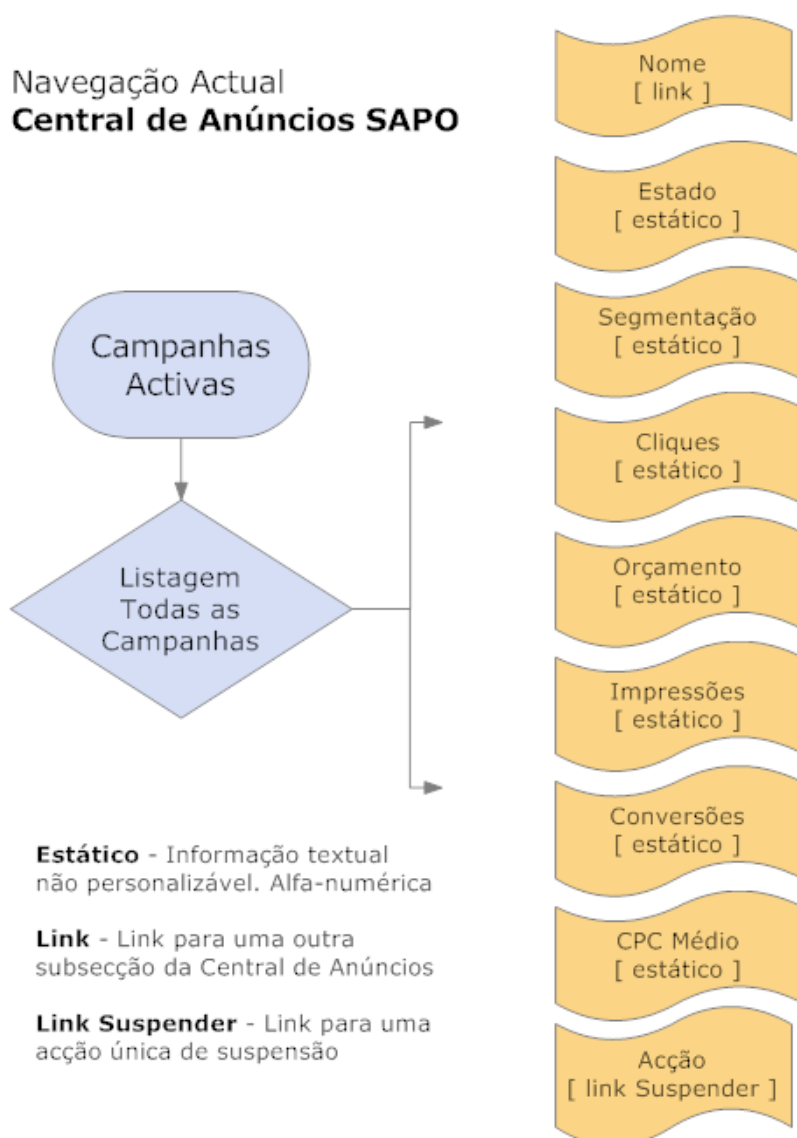


Figura 20: Esquema Actual de Navegação da secção Campanhas

Esta é a principal secção da Central de Anúncios a receber a maior actualização. Pretende-se implementar um sistema gráfico sempre que é apresentada informação estática. Esta secção, com uma implementação gráfica competente, poderá até a longo-prazo substituir a secção de Relatórios. Isto é, aprofundando as opções de visualização de campanhas e aplicando opções de exportação e recorrência, teremos as mesmas funcionalidades que a secção “Relatórios”.

Podemos verificar que colunas que, quer por tendências como alterações de metodologias de edição/consulta, deixam de fazer sentido na actualidade. Um exemplo disso, a apresentação de uma coluna que demonstra o “Estado” das campanhas, quando

a listagem é apenas de Campanhas Activas. Essa coluna é redundante pois será apresentada, supostamente, sempre com o estado “Activo”.

Assim, propõe-se a eliminação destas e outras colunas e a sua substituição por gráficos. Esses mesmos gráficos deverão complementar as informações apresentadas de forma estática e possibilitar uma interactividade.

Para além disso, propõe-se a implementação de um pequeno formulário, de importância vital, que permita ao utilizar definir períodos de consulta. Deverá ser apresentada a opção de uma data de início e data fim. Poderá ser complementado com as opções de “Hoje”, “Ontem”, “Últimos 7 dias”, “Último Mês” ou personalizar com precisão Data Início/Data Fim, onde D corresponde ao dia actual.

<i>Textual</i>	<i>Correspondência</i>
Hoje	D a D
Ontem	D-1 a D-1
Últimos 7 dias	D-7 a D-1
Semana passada	[penúltimo Domingo] a [último Domingo]
Este Mês	[1º dia do mês actual] a D
Mês passado	[1º dia mês anterior] a [Último dia mês anterior]
Últimos 90 Dias	D-90 a D
Sempre	[abertura de conta] a D
Data Personalizada	Inquirir Data Início e Data de Fim

Tabela 5: Estrutura Proposta Metas Temporais

Navegação Proposta Central de Anúncios SAPO

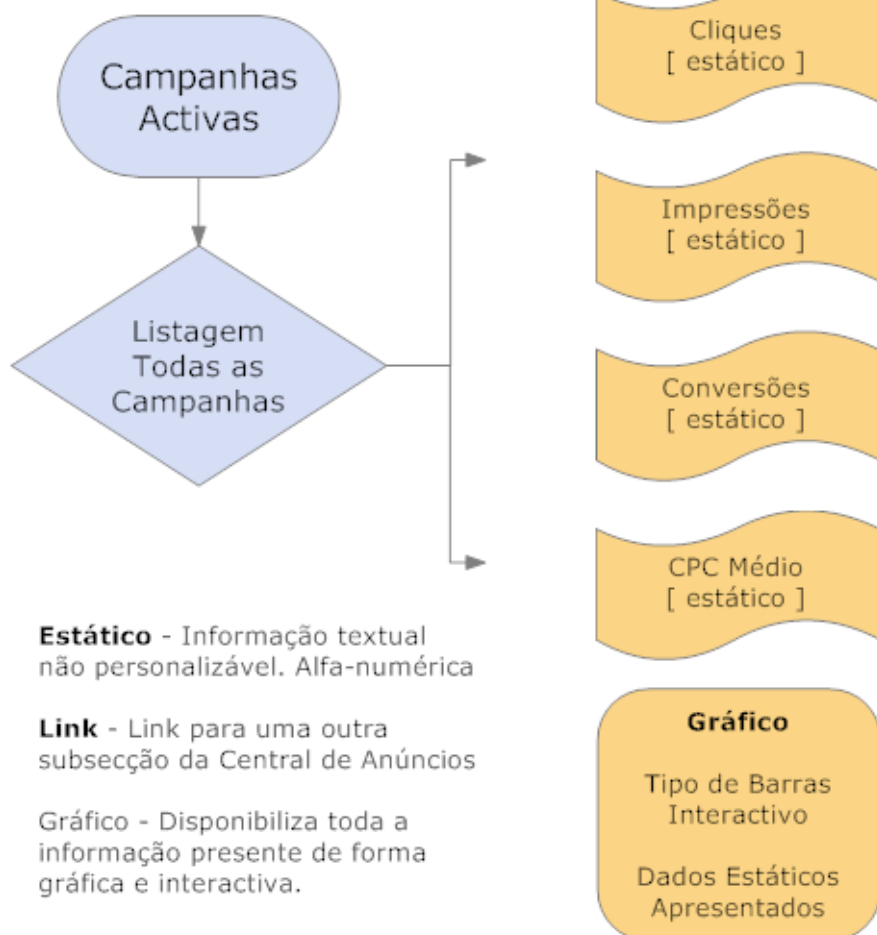


Figura 21: Esquema Sugerido de Navegação da secção Campanhas

5.7. Necessidade de Processos de Cache

Um sistema de cache permite registar em memória temporária, dados compilados, que mais tarde a serem consultados, são directamente acedidos sem nova computação. Garante-se com um sistema de cache uma rapidez na obtenção de dados pre-compilados, e ao mesmo tempo salvaguardam-se recursos dos servidores. Existem diversos mecanismos de cache assim como frameworks tanto em PHP como ao nível do

servidor. No caso, face à especificação dos elementos, optou-se por criar um sistema de cache com base no sistema de “registo de ficheiros para armazenamento de dados”.
[G06F13/00]

5.7.1 Introdução e Motivação

Qualquer anunciante que utilize a Central de Anuncios do SAPO, irá exigir rapidez na apresentação dos relatórios. Entende-se um anunciante como um utilizador experiente da internet, que a utiliza no seu dia-a-dia. Assim, deverá ser garantido uma robustez na apresentação dos resultados.

Partindo do principio que a informação passada não poderá ser nunca alterada, isto é, as impressões, cliques e conversões (seja a de ontem seja a de uma semana atrás), serão as mesmas quer sejam consultadas hoje ou daqui a um mês, poderemos ir registando em ficheiros de cache sempre que esses dados são acedidos pois são estáticos. Quando existe um pedido de relatório, o primeiro passo a efectuar é verificar se esses mesmos dados já foram requisitados. Se sim, acede-se directamente à cache. Em caso contrário, acede-se à base de dados, compila-se os dados, regista-se na cache e apresentam-se depois ao anunciante. Podemos interpretar o tempo de registo/leitura de cache quase nulo, não influenciando na experiência de navegação do anunciante.

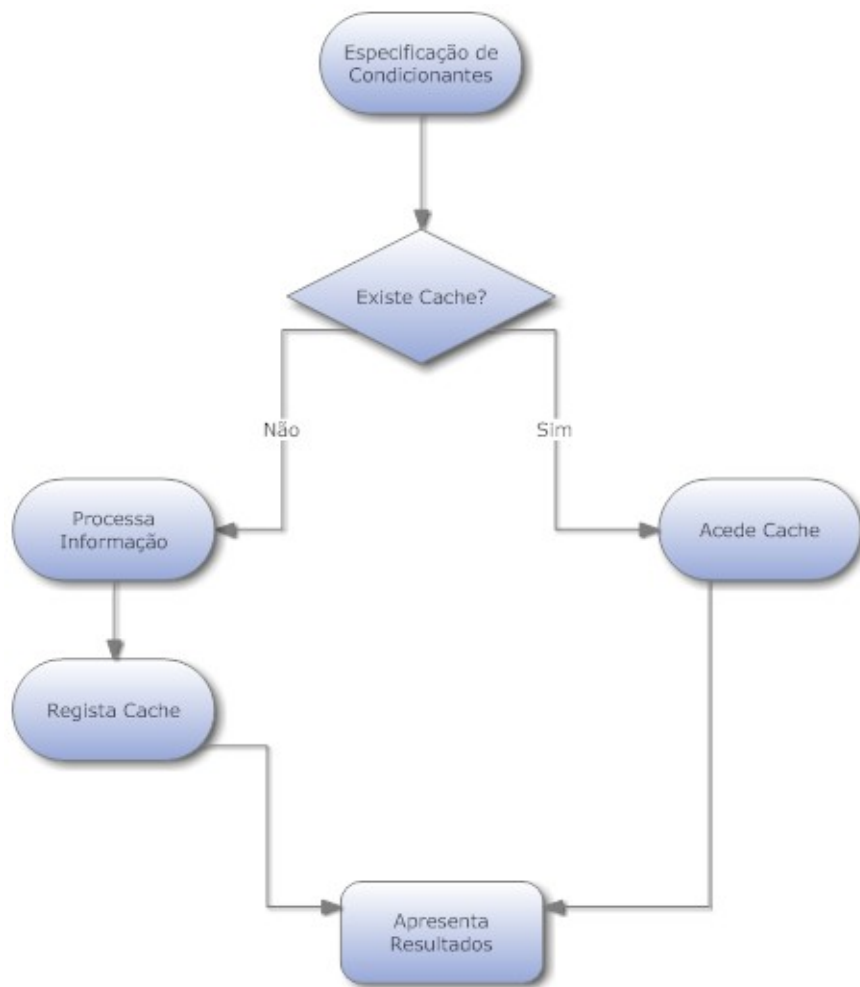


Figura 22: Sistema de Cache

5.7.2 Resultados Obtidos

Efectuou-se a medição de tempo da execução dos scripts em execução, com a utilização de uma função desenvolvida para o efeito. Assim, começou por se definir uma função que devolva o tempo actual em *timestamp*.

Depois, no início da execução do script, efectua-se o chamamento da função respectiva.

A partir deste momento, o tempo está a contar. No final do script, volta-se a consultar o *timestamp*, e assim, o tempo de execução corresponde à diferença de tempo entre o início da execução e o final da execução.

Assim, após implementação de algoritmos de medição de teste, efectuou-se algumas medições as quais aqui se registam:

<i>Registo de Tempo de Execução – Gráfico Impressões+Cliques+Conversões</i>			
<i>Teste</i>	<i>Cache Inactiva (s)</i>	<i>Cache Activa (s)</i>	<i>Variação</i>
1/11, Impressões	0,19	0,01	-0,18
11 a 13/1, Imp.Cliques e Conversões	2,63	0,01	-2,62

Tabela 6: Tabela de Registo de Tempos de Execução cache/sem cache no Gráfico #1

<i>Registo de Tempo de Execução – Gráfico CTR+Conv Rate + Impression Share</i>			
<i>Teste</i>	<i>Cache Inactiva (s)</i>	<i>Cache Activa (s)</i>	<i>Variação</i>
1/11 a 13/1, CTR	14,9	0,01	-14,8
1/11 a 31/12, Conv Rate e Imp.Share	16,13	0,01	-16,12

Tabela 7: Tabela de Registo de Tempos de Execução cache/sem cache no Gráfico #2

<i>Registo de Tempo de Execução – Gráfico Adg ID + Wrd ID</i>			
<i>Teste</i>	<i>Cache Inactiva (s)</i>	<i>Cache Activa (s)</i>	<i>Variação</i>
Adg ID 115300 Limites 10	0,29	0,01	-0,28
Ad ID 139576 Limites 15 (Pie Chart)	0,2	0,01	-0,19
Ad ID 140778 Limites 10 (Pie Chart)	0,18	0,01	-0,17

Tabela 8: Tabela de Registo de Tempos de Execução cache/sem cache no Gráfico #3

Pela análise das tabelas e os resultados obtidos, pode ser concluído com todo o rigor exigido, que o sistema de cache é funcional e permite um acesso constante de uma centésima de segundo, independentemente do esforço computacional efectuado ao gerar na primeira vez. Em gráficos especialmente complexos, como o gráfico CTR, Conversion Rate e Impression Share em que a média ronda os 15 segundos, será imperativo implementar tal sistema.

6. Conclusão e Trabalhos Futuros

Na realização deste projecto, face a algumas das limitações identificadas, muitas delas que fogem até dos seus objectivos, foram identificadas etapas a uma possível implementação. Assim, tendo em conta que um conceito/novidade na questão de anúncios rapidamente evolui, geralmente num prazo de 6 meses fica desactualizado, será fulcral uma constante actualização do produto final. Novas versões das bibliotecas aqui usadas estão sempre a ser lançadas, com inovações, algoritmos mais eficientes e erros corrigidos.

Assim, identificam-se as áreas da Central de Anúncios que deverá estar em constante evolução.

- **Plano Técnico**

Face às inúmeras utilizações de frameworks e ferramentas de terceiros, deverá ser garantida que cada uma delas se encontra funcional e actualizada. O correcto será actualizar cada ferramenta assim que estiver disponível uma nova versão e substituir, ou até remover por completo, por motivos de segurança, sempre que se verifique que foi descontinuada.(Tabela 76)

- **Algoritmos/Contextualização**

Uma das lacunas em termos de capacidades do projecto que foram identificadas, foi a limitação das campanhas no motor de busca do SAPO e em certos sites/parceiros. O SAPO

possibilita ao anunciante com todo o controlo, definir onde deseja aparecer os seus anúncios, quer em termos de pesquisas quer em termos de páginas.

Assim, é oferecida ao anunciante a possibilidade de anunciar por “carros”, mas muito embora, empiricamente seja reconhecido o sector que quer anunciar, se não inserir directamente a palavra-chave “automóveis”, os seus anúncios não serão apresentados. Entende-se por isso que será fulcral desenvolver e implementar um sistema de linguagem natural que possibilite aos anunciantes efectuar licitações em palavras não inseridas como sinónimos, erros-ortográficos e outras equivalentes. Uma possível resolução seria a oportunidade do anunciante licitar por “temas” ou “conceitos” ou “tópicos”¹⁶ em vez de palavras-chave. Um anunciante que licitasse pelo tema “automóvel” seria o equivalente a licitar por “carros” ou “automóvel” ou “veículos motorizados”. Da mesma forma, o anunciante ao invés de licitar páginas ou sites, licitaria no contexto das páginas. Grande maioria dos sites são dinâmicos e embora um anunciante do sector automóvel licite num jornal por este mencionar um grande artigo de carros, se o anunciante não estiver atento, quando esse mesmo artigo for substituído por outro que nada tem a ver, continuará a licitar em conteúdo nada relevante. Assim, recomenda-se que a evolução em termos de disponibilidade de anúncios seja no sentido da **linguagem natural**[NLSSA] e **contextualização em tempo real** como inovou o motor de busca Google.[GLAS][GL]

O SAPO regista informação sobre o historial de navegação dos seus utilizadores. Assim, com essa informação poderá ser mantido um histórico individualizado mediante os padrões de navegação e podemos assim indentificar padrões de utilização em diversos tópicos tanto individualizados como globais. Neste sentido, podemos apresentar ao utilizador anúncios mais relevantes pois foram identificados padrões e interesses. Só será permitido em grandes sites de tráfego como o SAPO e sempre com o princípio de garantir que a violação da privacidade do utilizador não é posta em causa.[MWU]

◦ **Conceitos e Negócio**

O modelo presente actualmente na plataforma de Anúncios SAPO, baseia-se apenas e só em campanhas em que o pagamento é por clique. Assim, face a futuras implementações de sistemas, um possível modelo a adoptar seria o pagamento ao invés de um clique num anúncio, o pagamento de uma determinada acção pré-definida no anunciante. Neste caso, um anunciante de seguros poderia remunerar o SAPO por cada contacto realizado e um comerciante de automóveis por cada carro vendido. Este é um modelo em evolução e elimina grande parte da fraude típica dos modelos Pay-Per-Click. Um caso actual de sucesso em tal modelo, é a plataforma de afiliação nacional Netlucro[NETLUC] onde assenta apenas no modelo de partilha de receitas entre a sua rede e as acções pré-definidos dos seus anunciantes.

16 Recentemente o programa Google Adwords disponibilizou a possibilidade de segmentar as campanhas por “Tópicos e Contextos” pré-definidos ao invés de sites ou por palavra-chave na sua rede de conteúdo. <http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=66761> [25/05/11]

Uma opção passível de ser igualmente executada e menos radical, seria a implementação de outros sistemas de licitação. Seja pela possibilidade de o anunciante licitar, por vezes, em que os seus anúncios são apresentados (desprezando os cliques – modelo CPM), ou então remunerar por cada conversão efectuada no seu site. Estas soluções entendem-se ser mais robustas mas necessitam de maior controlo por parte da plataforma.

Uma vez que os sistemas de *tracking*¹⁷ estão em constante desenvolvimento e as questões de alegadas violações de privacidades estão sendo culmatadas com legislação específica, é possível utilizar de momento o historial de navegação de um utilizador que corresponde ao tipo de páginas vistas, quando foram vistas e de que maneira, recomendar serviços e produtos específicos. Como exemplo, um utilizador que clique em anúncios de meias, sabemos que alegadamente consumiu um produto que irá provavelmente voltar a consumir passado certo tempo. Então a esse utilizador serão apresentados anúncios de meias passado um mês. Esta metodologia tem um impacto muito grande no melhoramento do ROI das campanhas.[INT59][MWEUBINT]

◦ **Análise concorrencial relativa**

Como em qualquer campanha, é natural a existência de um anunciante concorrente a oferecer os mesmos produtos e serviços. Será pertinente para um determinado anunciante tomar conhecimento da performance das campanhas em relação à concorrência.

Assim, disponibiliza-se uma ferramenta que permite examinar a actividade de uma campanha nas últimas semanas e apresentar categorias que representam o target conhecido. As categorias correspondem aos termos de pesquisa, textos de anuncios e páginas de destino, que serão depois categorizados. Para cada categoria identificada, é apresentado um gráfico de barras que compara a performance individual da campanha com a média de performance de outros anunciantes identificados nessas mesmas categorias.

Com esta possibilidade, é disponibilizado ao anunciante uma visão geral da campanha em relação a outras campanhas, e compará-las em termos de CTR, reach de impressões, cliques e posições médias. A informação é anónima e não será possível identificar individualmente os concorrentes, baseando-se sempre no global dos dados.

¹⁷ Tracking corresponde aos mecanismos tecnológicos que permitem registar com sucesso acontecimentos em páginas web, tais como cliques, visualizações, independentemente do tempo, estando limitados às durações do *cookie*.

Em termos de Análise concorrencial relativa, destacamos o sistema do Google Adwords[AGLWEBR] que possibilita aos anunciantes uma comparação com outros anunciantes similares.

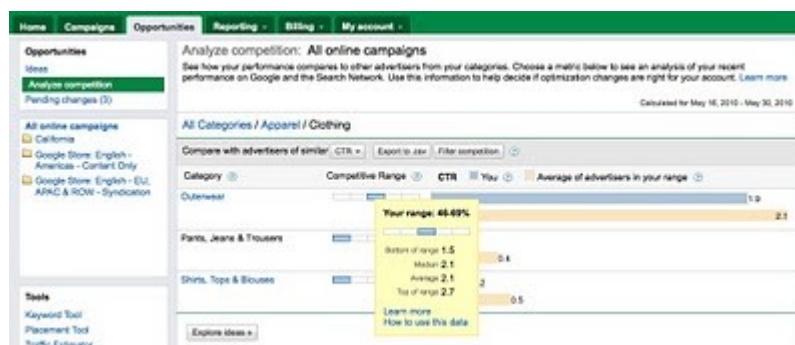


Figura 23: Google Adwords - Análise concorrencial relativa

Existem também ferramentas tanto em software como web-based que permite a análise da concorrência. É certo que se perde a parcialidade ao ser um terceiro a fornecer os dados, mas não é possível garantir que esses mesmos dados são exactos e íntegros pois em regra, as plataformas de anúncios não se relacionam com tais soluções por interesses de negócio e tecnológicos.¹⁸

◦ Outros

Algo fundamental em qualquer sistema que exija pagamentos em sistema de pré-pagamento ou a diferido, é a implementação de soluções de pagamento. Assim, entende-se ser crucial manter os últimos meios de pagamento e respectivo processamento automático. Pagamento de Compras (Multibanco), Cartão de Crédito (VISA/MASTERCARD) e Pay Pal (e-wallet) constituem o mínimo essencial a ter presente em qualquer negócio B2B.[PAGPPC]

Outro grande desafio de qualquer programa de anúncios, é o combate ao SPAM. Tal como a Symantec[SYMSPAM] constatou no seu estudo, os novos spammers utilizam plataformas como o Google Adwords para efectuar phishing e outras técnicas maliciosas. A certos níveis, o combate à fraude e ao spam, poderá ser o que ditará a sobrevivência de muitos programas de anúncios.

¹⁸ A Microsys tem uma solução que permite a avaliação de concorrência com dados próprios: <http://www.microsystools.com/products/keyword-research/help/keyword-research-competition-analysis/> [26/05/2011]

6.1. Conclusões

Para obter uma melhor gestão e optimização das campanhas dos anunciantes, deverá existir uma gestão central, apoiada por um departamento responsável com competências técnicas e de negócio (Central de Anúncios). Para tal, deverá ser seguido um modelo que defina as regras de consulta, gestão e optimização de campanhas, orientado para os seus objectivos. Regra geral, esses objectivos definem-se por um aumento do Retorno de Investimento (ROI).

De acordo com o problema inicialmente proposto¹⁹, foram identificadas lacunas na gestão de Anúncios SAPO e respectivas ferramentas, devido à omissão de procedimentos adequados quer de optimização de campanhas, quer de visualização de relatórios. No presente caso, como é de importância vital ao anunciante ser dado a conhecer os seus relatórios (em forma de estatísticas temporais (ou não)), verificou-se a inexistência de qualquer método de visualização gráfico. A informação é estática e sem interactividade.

Assim, no passo inicial, procedeu-se a um levantamento do estado actual²⁰ a nível de disponibilização de informação, interacção oferecida e ferramentas de optimização. O mesmo foi feito de uma forma geral, ao tipo de dados processados e a maneira como estes são pré-compilados.

No estado actual, verificou-se que a alteração de alguns procedimentos visuais, se tornaram necessários para disponibilizar informação de uma maneira mais rápida, visualmente mais perceptível e com interacção. Com este procedimento, o SAPO passou a oferecer gráficos nos mais diversos tipos, e o anunciante passou a poder interagir e manipular os dados visualizados instantaneamente.

Por outro lado, encontrava-se em falta um sistema que auxiliasse o anunciante na optimização das mesmas. Como a optimização consiste em disponibilizar sugestões automáticas, de potenciais alterações às campanhas com melhores resultados futuros, estamos perante sugestões e decisões de negócio que contém algum risco. Essas optimizações serão implementadas exclusivamente pelo anunciante, limitando-se o sistema apenas a sugeri-las, mediante algoritmos complexos que optimizem a conversion rate, o CTR dos anúncios, a refinação do orçamento e até lista de palavras.

A nível de formas de visualização, vários modelos foram abordados. O modelo de visualização gráfica *Polaris*, consistem numa framework em JavaScript que permite

¹⁹ Descrição do Problema (capítulo 1.2)

²⁰ Caracterização do Estado Actual (capítulo 2.4.2)

integrar dados estáticos em gráficos SVG. O modelo *InfoVis*, mais adequado para grandes quantidades de dados mas baseado em Java, é uma boa opção para análises individuais de n-variáveis. Não é necessariamente uma ferramenta web-friendly²¹ mas graças à sua robustez, sendo desenvolvida em Java, foi considerada para o efeito. O *Protovis*, uma frameworks em JavaScript, que tal como o *Polaris* permite a interactividade com o utilizador, mas com uma modularidade suficiente que permite separar ao programador os dados propriamente ditos e os gráficos.

Conclui-se, assim, que o modelo de visualização mais ajustado é o *Protovis*, dado que estamos perante um sistema complexo, que requer modularidade e independência entre dados/visualização.

A nível de formas de optimização, foram estudados vários algoritmos para o efeito. Foi proposto um sistema de optimização de palavras-chave (linguagem natural, sinónimos, capitalização e plurais), assim como previsões de CTR de novos anúncios e optimizações de lances de CTR. Para além disso, foi estudada uma possibilidade de optimização de orçamentos (*budget*), com o objectivo de melhor atribuir os orçamentos, quando os mesmos são limitados, a recursos mais importantes dos objectivos da campanha. As optimizações são meramente informativas, pelo que houve a necessidade de criar um módulo novo que disponibilizasse as ferramentas e apresentasse os seus resultados, com o fornecimento de poder de decisão ao anunciante: o anunciante, poderá após consultar as optimizações sugeridas, aplicar ou não as mesmas nas suas campanhas online.

Logo após a definição dos modelos a adoptar, procedeu-se às suas implementações, sendo estas efectuadas paralelamente a uma refinação de navegação da Central de Anúncios SAPO, necessária para adoptar as soluções apresentadas. Foi assim realizado um levantamento do esquema de navegação, e adaptado o modelo, actualizando-o removendo informações redundantes e implementado as novas ferramentas nomeadamente o estilo gráfico e interactivo oferecido pelos modelos de visualização.

Sendo o universo de modelos gráficos a aplicar bastante diverso e dados novos estão a ser constantemente atribuídos, não foi possível implementar todas as combinações possíveis de gráficos, muito embora sejam aqui apresentados. Relativamente às optimizações, entende-se como um grande desafio que actualmente define a qualidade das centrais de anúncios existentes, pelo que a sua implementação deverá ser efectuada num departamento a criar especificamente para o efeito, onde seja aliado o poder dos algoritmos aqui apresentados, com o estudo humano muitas vezes insubstituível.

Assim, muito embora tenham sido identificadas lacunas e implementadas soluções, existe um trabalho continuo a ser efectuado que deverá ser encarado como uma evolução permanente na Central de Anúncios. Caso contrário, estaremos rapidamente

21 Entenda-se *web-friendly*, uma solução desenvolvida para os *browsers* de internet.

perante um sistema obsoleto. As abordagens técnicas e teórica são assim aqui apresentadas, com o objectivo de partir para a evolução necessária.

7. Referências e Bibliografia

- [ADRMTC] Aranyak Mehta Amin Saberi Umesh Vazirani Vijay Vazirani, AdWords and Generalized On-line Matching,
- [AGILE2d] Human Computer Interaction LAB, ,
<http://www.cs.umd.edu/hcil/agile2d/index.shtml>
- [AGLWEBR] Webrageous, Adwords Analyze Competition, 2010
- [AXIIS] AXIIS, , <http://thejit.org/docs/>
- [BING] Bing!, , <http://www.bing.pt>
- [BINGSAPO] Pesquisas no BING com Anuncios SAPO, ,
http://ajuda.sapo.pt/faq.html?faq_id=35508
- [CASEST] ReStockIt.com, Supply and Demand,
- [CASETR] Google Adwords, Travel Industry Knowledge Center ,
- G06F13/00: APOSTOLIDES VASILIS J [US]; BERTSCHI JASON [US]; BOYLES JOHN R [US]; HOSTENY JOSEPH NEVI IV [US]; KAZAR MICHAEL L [US], FILE STORAGE SYSTEM, CACHE APPLIANCE, AND METHOD ,
- [GACR2] Jordan McCollum , Google Says Average AdWords CTR at 2%, 2010
- [GGICASE] GOOGLE, Google AdWords Case Study – Improving Landing Page Quality,
- [GL] Google Expands Advertising Monetization Program for Websites, ,
<http://www.google.com/press/pressrel/adsense.html>
- [GLAS] Google Acquires Applied Semantics, ,
<http://www.google.com/press/pressrel/applied.html>
- [INFOVS] Jean-Daniel Fekete1, The InfoVis Toolkit,

- [INT59] Smith-Grieco, Anthony Nicoll, The Internet as recommendation engine : implications of online behavioral targeting, 2010
- [ISLOST] How to Avoid Lost Impression Share Due to Budget, , <http://searchenginewatch.com/article/2066276/How-to-Avoid-Lost-Impression-Share-Due-to-Budget>
- JGPPC: , PPC search engine marketing placement - the best way to advertise, 2010
- [MPORTES] M. Porter, An algorithm for suffix stripping,
- [MWEUBINT] Wu, H.; Gordon, M.; DeMaagd, K.; Fan, W, Mining Web navigations for intelligence, 2006
- [MWU] Harris Wua, Michael Gordonb, Kurtis DeMaagdb and Weiguo Fanc, Mining web navigations for intelligence, 2006
- [NETLUC] Netlucro, , <http://www.netlucro.com>
- [NLSSA] Kevin Bartz, Kevin Bartz and Kevin Bartz, Natural Language Generation for Sponsored-SearchAdvertisements, 2008
- [ONLmP] MARAMREDDY, SUSHMA, Online Matching Problem with Application: Google Adwords Problem, 2006
- [OPTKenv] How Can Google AdWords Optimization Help in Internet Marketing?, , <http://www.pubarticles.com/article-how-can-google-adwords-optimization-help-in-internet-marketing--1305005471.html>
- [PAGPPC] MasterCard Advisors, The State ofB2B Payment Transformation, 2008
- [PREDICTCTR] Krzysztof Dembczyński, Wojciech Kotłowski, Wojciech Kotłowski, Predicting Ads' ClickThroughRate with Decision Rules,
- [PRFING] Catherine S. Panaguiton, PROFITING FROM PAY-PER-CLICK ADVERTISING MUST COME WITH A PRICE- A HEFTY ONE, AT THAT,
- [PROMBJH] Michael Bostock AND Jeffrey Heer, Protovis: A Graphical Toolkit for Visualization, 2009
- [PUBL4] Steve O'Keefe, Publicity on the Internet: Creating Successful Publicity Campaigns on the Internet and the Commercial Online Services, 1996
- [SYMSPAM] Symantec, Symantec: Spammers using AdWords , 2008
- [WEBSITEOPT] Andre B.King, Website Optimization, 2009
- [WWWPROT] Protovis, , <http://vis.stanford.edu/protovis/>

Anexo A: Anexos

I. Protovis: Exemplo de Aplicação

Um exemplo de código em Javascript utilizando PROTIVIS poderá ser:

```
// Cria painel 150x150
var vis = new pv.Panel()
    .width(150)
    .height(150);

// Adiciona réguas
vis.add(pv.Rule)
    .data(pv.range(0, 2, 1/2))
    .bottom(function(d) d * 80 + 1)
    .add(pv.Label);

//Adiciona as barras do gráfico
vis.add(pv.Bar)
    .data([1, 1.2, 1.7, 1.5, .7])
    .width(20)
    .height(function(d) 80 * d)
    .bottom(0)
    .left(function() this.index * 25 + 25)
    .anchor("bottom").add(pv.Label);

// Processa tudo
vis.render();
```

O exemplo acima obteremos:

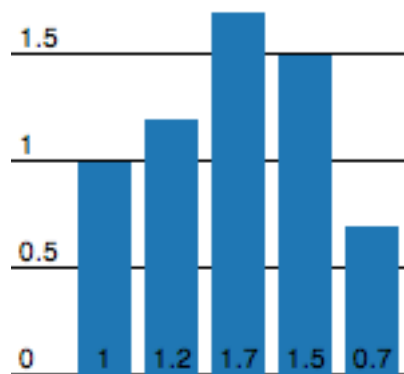


Figura 24: Exemplo de um gráfico/barras Protovis

II. Frameworks/Ferramentas Instaladas

Nome	Versão Instalada	URL
PHP		
Jquery		
Protovis		
MySQL		
phpMyAdmin		

Table 9: Ferramentas Utilizadas

II. Tabelas de dados SQL

Field	Type	Null	Default	Comments
adg_wrd_id	int(11)	Yes	NULL	
adg_id	int(11)	Yes	0	
wrd_id	int(11)	Yes	0	
adg_wrd_imp	bigint(20)	Yes	0	
ressions				
adg_wrd_clic	int(11)	Yes	0	
ks				
adg_wrd_cpc	float(9,2)	Yes	0.00	
adg_wrd_cre	datetime	Yes	NULL	
ated				

modified	adg_wrd_mo	timestamp	Yes	CURRENT_
	adg_wrd_del	tinyint(4)	Yes	TIMESTAMP
deleted				0
	usr_id	int(11)	Yes	NULL
	adg_wrd_url	mediumtext	Yes	NULL
	adg_wrd_del	int(11)	Yes	0
deleted_by				
	geo_id	int(11)	Yes	0
	local_tag	int(11)	Yes	0
	site_id	int(11)	Yes	NULL
in	adg_wrd_orig	smallint(2)	Yes	0
	sasi	double	Yes	0

Tabela 10: Tabela SQL adgroup_word

	Field	Type	Null	Default	Comments
	adg_id	int(11)	Yes	NULL	
	usr_id	int(11)	Yes	NULL	
	cam_id	int(11)	Yes	NULL	
	adg_descripti	varchar(120)	Yes	NULL	
on					
	adg_max_cpc	float(9,2)	Yes	NULL	
	adg_status	tinyint(3)	Yes	NULL	
	adg_deleted	tinyint(1)	Yes	0	
	adg_created	datetime	Yes	NULL	
	adg_modified	timestamp	Yes	CURRENT_	
url				TIMESTAMP	
	adg_default_	mediumtext	Yes	NULL	
	adg_medium	tinyint(1)	Yes	0	

Tabela 11: Tabela SQL adg_adgroups

Field	Type	Null	Default	Comments
site_id	int(11)	Yes	NULL	
affiliate_id	int(11)	Yes	NULL	
cluster_id	int(11)	Yes	1	
impressions	bigint(20)	Yes	0	
clicks	int(11)	Yes	0	
election_plug	varchar(80)	Yes	NULL	
in_description				
site_main_url	varchar(255)	Yes	NULL	
active	tinyint(1)	Yes	1	
is_mobile	tinyint(1)	Yes	0	
update_ts	timestamp	Yes	CURRENT_	
			TIMESTAMP	

Tabela 12: Tabela SQL Adsense_AffiliateSite

Field	Type	Null	Default	Comments
position_id	int(11)	Yes	NULL	
deal_id	int(11)	Yes	NULL	
template_id	int(11)	Yes	NULL	
palette_id	int(11)	Yes	NULL	
affiliate_id	int(11)	Yes	NULL	
site_id	int(11)	Yes	NULL	
impressions	bigint(20)	Yes	0	
clicks	int(11)	Yes	0	
last_update	datetime	Yes	NULL	
name	varchar(255)	Yes	NULL	
location_id	int(11)	Yes	NULL	

Tabela 13: Tabela SQL Adsense_PositionInSite

	Field	Type	Null	Default	Comments
	ads_id	int(11)	Yes	NULL	
	usr_id	int(11)	Yes	NULL	
	adg_id	int(11)	Yes	NULL	
	ads_title	varchar(255)	Yes	NULL	
	ads_line1	varchar(255)	Yes	NULL	
	ads_line2	varchar(255)	Yes	NULL	
	ads_url_redir	mediumtext	Yes	NULL	
ect					
	ads_url_displ	varchar(255)	Yes	NULL	
ay					
	ads_deleted	tinyint(1)	Yes	0	
	ads_created	datetime	Yes	NULL	
	ads_modified	timestamp	Yes	CURRENT_	
				TIMESTAMP	
	img_path	varchar(255)	Yes	NULL	
	img_title	varchar(50)	Yes	NULL	
	ads_impressi	bigint(20)	Yes	0	
ons					
	ads_clicks	int(11)	Yes	0	
	ads_type	tinyint(4)	Yes	0	

Tabela 14: Tabela SQL ads_ads

	Field	Type	Null	Default	Comments
	imp_date	date	Yes	NULL	
	bid_id	int(10)	Yes	NULL	
	ad_id	int(10)	Yes	NULL	
	count	int(10)	Yes	NULL	
	cluster_id	tinyint(3)	Yes	NULL	

Tabela 15: Tabela SQL bid_impressions

	Field	Type	Null	Default	Comments
	cam_id	int(11)	Yes	NULL	
	usr_id	int(11)	Yes	NULL	
	cam_descripti	varchar(120)	Yes	NULL	
on	cam_max_bu	float(9,2)	Yes	NULL	
dget	cam_start_dat	datetime	Yes	NULL	
e	cam_end_dat	datetime	Yes	NULL	
e	cam_status	tinyint(3)	Yes	NULL	
	cam_deleted	tinyint(1)	Yes	0	
	cam_created	datetime	Yes	NULL	
	cam_modifie	timestamp	Yes	CURRENT_	
d	cam_daily_co	float(12,2)	Yes	TIMESTAMP 0.00	
sts	geo_id	int(11)	Yes	0	
	target_site	tinyint(1)	Yes	0	
	site_targeted	tinyint(1)	Yes	0	
	content_netw	tinyint(1)	Yes	1	
ork					

Tabela 16: Tabela SQL cam_campaigns

	Field	Type	Null	Default	Comments
	click_id	int(10)	Yes	NULL	
	click_date	date	Yes	NULL	
	click_time	time	Yes	NULL	
	bid_id	int(11)	Yes	NULL	
	ad_id	int(11)	Yes	NULL	
	site_id	int(10)	Yes	NULL	
	position_id	int(10)	Yes	NULL	
	cost	int(10)	Yes	NULL	

Tabela 17: Tabela SQL clicks

Field	Type	Null	Default	Comments
conv_date	date	Yes	NULL	
conv_time	time	Yes	NULL	
ad_id	int(11)	Yes	NULL	
bid_id	int(11)	Yes	NULL	
count	int(10)	Yes	NULL	

Tabela 18: Tabela SQL conversions

Field	Type	Null	Default	Comments
imp_date	date	Yes	NULL	
site_id	int(10)	Yes	NULL	
position_id	int(10)	Yes	NULL	
ad_type	tinyint(3)	Yes	NULL	
count	int(10)	Yes	NULL	

Tabela 19: Tabela SQL site_impressions

Field	Type	Null	Default	Comments
usr_id	int(11)	Yes	0	
data	date	Yes	0000-00-00	
campanha	varchar(255)	Yes		
grupo	varchar(255)	Yes		
palavra	varchar(255)	Yes		
url	varchar(255)	Yes		
impressoes	int(11)	Yes	0	
clicks	int(11)	Yes	0	
custo	float(7,2)	Yes	0.00	
posicao_medi	float(7,2)	Yes	0.00	
a				
ctr	float(7,2)	Yes	NULL	
conv_rate	float(7,2)	Yes	NULL	
impression_s	float(7,2)	Yes	NULL	
hare				
cam_id	int(11)	Yes	0	
adg_id	int(11)	Yes	0	
wrd_id	int(11)	Yes	0	
adg_wrd_id	int(11)	Yes	0	
ads_id	int(11)	Yes	0	
conversions	int(10)	Yes	0	

Tabela 20: Tabela SQL usr_impressions_report

Field	Type	Null	Default	Comments
wrd_id	int(11)	Yes	NULL	
adg_id	int(11)	Yes	NULL	
impressoes	int(10)	Yes	0	
clicks	int(10)	Yes	0	
ctr	float(7,2)	Yes	0.00	
conversions	int(10)	Yes	0	
custo	float(7,2)	Yes	0.00	
avg_posicao_	float(7,2)	Yes	0.00	
media				
imp_share	float(7,4)	Yes	NULL	
click_share	float(7,4)	Yes	NULL	
conv_share	float(7,4)	Yes	NULL	
custo_share	float(7,4)	Yes	NULL	

Tabela 21: Tabela SQL wrd_adg_performance_report

Field	Type	Null	Default	Comments
wrd_id	int(11)	Yes	NULL	
impressoes	int(10)	Yes	0	
clicks	int(10)	Yes	0	
ctr	float(7,2)	Yes	0.00	
conversions	int(10)	Yes	0	
custo	float(7,2)	Yes	0.00	
avg_posicao_	float(7,2)	Yes	0.00	
media				

Tabela 22: Tabela SQL wrd_performance_report

Índice Remissivo

C

call-to-action, 27

CPA, 58

cpc, 75, 76

CPC, 8, 11, 20, 21, 24, 25, 27, 55-57, 58

CTR, 67

G

Google Adwords, 68

J

Javascript, 37, 38, 47-49, 74

JavaScript, 16, 69, 70

jQUERY, 16

Jquery, 75

JQUERY, 48

P

Pay-Per-Click, 66

R

ROI, 15, 17, 26, 27, 29, 58